ED 043 356

CALIF. STATE LIBRARY. PROCESSING CENTER DESIGN AND SPECIFICATIONS, VOL. V., COST ANALYSIS. SUPPLEMENTAL VOLUME. HARGROVE, THOMAS L., STIRLING, KEITH H. CALIF. UNIV., BERKELEY. INST. OF LIBRARY RESEARCH. 70. 94P.

DOCUMENT RESUME

ED 043 356 LI 002 148

AUTHOR	Hargrove, Thomas L; Stirling, Keith H.
TITLE	California State Library: Processing Center Design and Specifications, Vol. V: Cost Analysis. Supplemental Volume
INSTITUTION	California Univ., Berkeley. Inst. of Library Research.
SPONS AGENCY	California State Library, Sacramento.
PUB DATE	70
NOTE	94p.
EDRS PRICE	EDRS Price MF-$0.50 HC-$4.80
DESCRIPTORS	*Automation, Computer Programs, *Cost Effectiveness, *Libraries, Library Services, *Library Technical Processes, Program Design, *State Libraries, Systems Development
IDENTIFIERS	California, *Library Automation

ABSTRACT

Presenting this cost analysis as a supplemental volume, separate from the main report, allows the chief activities in implementing the Processing Center Design to be correlated with costs as of a particular date and according to varying rates of production. In considering the total budget, three main areas are distinguished: (1) Systems Development, (2) Conversion and Maintenance, and (3) Output Services. Separate costs are stated for the different operational phases of Monographs and Serials, although both types of data are considered as sharing management and development costs. The analysis first covers the purpose, scope, allocation and descriptive basis of the three main areas for a six-month preparatory period to establish the Center and three years full production. Appendices include a reprint of the preliminary organizational design (with minor revisions) and supporting details for the three main areas. Volumes I-IV of this report are available as ED number 036 305 to 036 308. (Author/NH)

EDO 43356

CALIFORNIA STATE LIBRARY:

PROCESSING CENTER DESIGN AND SPECIFICATIONS

Vol. V

COST ANALYSIS

INSTITUTE OF LIBRARY RESEARCH
UNIVERSITY OF CALIFORNIA, BERKELEY

July 1970

Z 671 .C2 no. 17

+6

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED043356

CALIFORNIA STATE LIBRARY:

PROCESSING CENTER DESIGN AND SPECIFICATIONS

VOL V: COST ANALYSIS

Supplemental Volume

1970

LT 002 148

by

Thomas L. Hargrove

and

Keith H. Stirling

INSTITUTE OF LIBRARY RESEARCH
UNIVERSITY OF CALIFORNIA, BERKELEY

F O R E W O R D

This cost analysis for implementing California State Library: Processing Center Design and Specifications is presented here as a supplemental fifth volume of the report. This separate format allows the chief activities in implementing the design to be correlated with costs as of a particular date and according to varying rates of production.

In considering the total budget, three main areas are distinguished: Systems Development, Conversion and Maintenance, and Output Services. Separate costs are stated for the different operational phases of Monographs and Serials, although both types of data are considered as sharing management and development costs.

The analysis first covers the purpose, scope, allocation, and descriptive basis of the three main areas, presented as an Introduction to the Summary Cost Estimates. These summaries are for a six month preparatory period to establish the Center, and three years full production covering:

A. Systems development of computer programs, training and research.

B. Establishment, maintenance, and expansion of Center (in terms of Personnel, Services, and Expenses).

C. Operational unit-costs of conversion and maintenance.

D. Output services, or typical costs to print book catalogs or selective lists.

Appendices follow which include a reprint of the preliminary organizational design (with minor revisions), and supporting details for the three main areas.

We should like to acknowledge especially the contribution of Dennis Fried and Ruth Dunham on Serials, and Philip Bellman on numbering catalog cards. We have also relied heavily on the results of previous work done by David Berg on Output Services.

We also want to express our gratitude to the office staff who helped so much in preparing these pages: Joan Chan, Linda Child, Kitty Colburn, Elizabeth Ford, Bettye Geer, Marion Gordon, Linda Herold, Linda Horton, Pamela Mitchell, Judi Sutliff, and Connie Torii.

The work in this report was done under the supervision of Don Sherman, the Project Manager for the California State Library Processing Center contract, with the advice and assistance of Project Coordinator Ralph M. Shoffner.

T.L.H.
K.H.S.

Berkeley, California

C O N T E N T S

F O R E W O R D i

C O N T E N T S iii

I N T R O D U C T I O N

A. Purpose of Costing Information 1
B. Scope of Estimates 1
C. Descriptive Basis of Estimates 5
1. Personnel 6
2. Services 7
3. Allocation 8
4. Systems Development. 9
5. Conversion and Maintenance 12
6. Output Services--Catalogs. 15

C O S T S U M M A R I E S 23

A. Systems Development of computer programs, training, and research 25
B. Establishment of Center, Conversion, and Maintenance 27
C. Unit Costs--Full Operation 32
D. Output Services (Printing Book Catalogs) 35

A P P E N D I C E S 43

I. Preliminary Organizational Design. 44
II. Systems Development. 51
Summary. 52
A. Programming. 52
B. Training 53
C. Research 53
D. Management 53

C O N T E N T S (cont.)

A P P E N D I C E S

III. Conversion and Maintenance 54
Summary. 55
A. Initial Conversion:
MONOGRAPHS
1. Production and Control 57
2. Editing. 62
3. Keying 64
4. Computer Processing. 66
5. Proofing 72
SERIALS
6. Complete Initial Conversion. 79

B. Conversion of other libraries. 81
C. Maintenance of established files (MARC). 86
D. Authority Verification 88
E. Data Retrieval 91

IV. Output Services Cost Factors 94

S E L E C T E D B I B L I O G R A P H Y 101

I N T R O D U C T I O N

A. PURPOSE OF COSTING INFORMATION

Costing is considered here as an essential and important part of the information needed by Management to achieve its goals in successive phases of computerizing library records. It can substantially assist in planning, budgeting, and implementing those goals.

In the first phase of such computerization, the research study Catalogs in Book Form, the purpose of costing information was well described in Hayes and Shoffner's Economics of Book Catalog Production, from which the equations for machine processing were taken (Catalogs, p. 46). Costing information

> provides "order of magnitude" estimates which will help many people decide whether they are interested at all. If there is interest, a small number of rates must be specially determined to obtain more accurate estimates.
>
> (Economics, p. 5/6)

This initial costing information has satisfied the interests of CSL that the cost range and feasibility of computer-based Book Catalogs is competitive on its own merits compared to other approaches, and can be done with an outside computer.

In the second phase presented here, design and specifications, costing information is considered a continuation and refinement of those initial estimates.

B. SCOPE OF ESTIMATES

Estimates are projected for a preparatory six months and three years full production in establishing, maintaining, and expanding the Center. Three broad divisions of activity describe the work for which Management is responsible during this period: 1. Systems Development, 2. Conversion and Maintenance, and 3. Output Services.

1. Systems Development covers the immediate computer programming, or software, to be done by the outside computer center, as well as training to be done inhouse or outside. It also includes the ongoing research, whether inhouse or outside, that is needed for applying both immediate and long range advances in the computer and manual aspects of operational systems. Management would coordinate these three aspects of programming, training, and research, and their application within the Center.

2. Conversion and Maintenance covers Management's initial preparation and overall supervision of initial and subsequent conversions and maintenance. It also includes the permanent staff's source-preparation and control of catalogs during conversions, and regular maintenance of files after conversions are completed. Most importantly in costs, it also includes the work by outside services and additional temporary personnel for editing, keying, computer processing, and proofing the huge retrospective catalogs and interim changes during conversion.

3. Output Services covers Management's planning and coordinating of the printing of the first local and Union book catalogs of monographs or serials, as well as any selective lists useful for library procedures or public access. Also included are the permanent staff's exception-coding and control of the file, plus various outside services needed for computer preparation of the file, through printing and distribution of the finished catalogs or lists.

All these Implementation activities of the Center might be summarized in chart form, for successive three month periods, on the following page.

I M P L E M E N T A T I O N --- C S L - P R O C E S S I N G C E N T E R

(Permanent Staff) (Outside/Additional Staff)

Quarterly Periods In Months	M A N A G E M E N T Coordinating Systems Development, and Operations:	C O N V E R S I O N / P R I N T I N G			
		Monographs		Serials	
		1st library	2nd & other libraries	1st library	2nd & other libraries
Preparatory					
First six months	Establish Center Prepare 1st Monograph Conversion				
Full Production					
1 - 3rd	Prepare 2nd Monograph Conversion	Initial Conversion			
4 - 6th	Prepare 1st Serial Conversion	Full Conversion	Initial Conversion		
7 - 9th	Prepare 2nd Serial Conversion		Full Conversion	Initial Conversion	
10 - 12th	Continuing			Full Conversion	Initial Conversion
13 - 15th					Full Conversion
16 - 18th					
19 - 21st	Supervision				
22 - 24th	Prepare Print of 1st Catalog				
25 - 27th	Maintain all post-conversion entries and printing	Print 1st Catalog			
28 - 30th		Print 1st Suppl.			
31 - 33rd	Prepare Print of Union Catalog	Print 2nd Suppl.			
34 - 36th	Maintain & Expand Center	Print 3rd Suppl.			
	Print Union Catalog				

C. DESCRIPTIVE BASIS OF ESTIMATES

(1) Personnel

(2) Services

are first described as basic functions underlying proposed plan of six months' preparatory period followed by initial three year's full production. The plan itself is then described briefly in terms of proposed

(3) Allocation of costs

(4) Systems development

(5) Conversion and maintenance

(6) Output services

These descriptions are a verbal introduction to Cost Summaries in the next section.

1. PERSONNEL

Permanent Personnel

A more detail description is given in the report, Vol. II, pp. 253/8, reproduced as one of the Appendices in this volume.

Center Coordinator:	Administrative head of Center, knowledgeable about both systems analysis and library procedures.
Operations Manager:	Operational head of Center, knowledgeable about, technical data processing and library procedures.
Data Supervisor:	Responsible for training, supervising, and coordinating editing and proofing for Monographs and Serials.
Technical Editor:	Responsible for documentation work of Center.
Secretary:	Responsible for secretarial work done by Center.
2 Clerks:	Responsible for clerical procedures in conversion and maintenance.
Additional Personnel:	(Assuming concurrent implementation of Monographs and Serials).

MONOGRAPHS

2 Senior Editor/ Proofers	Responsible for supervising and coordinating editors/proofers.
8 Editors/Proofers:	Responsible for coding and editing sheets, proofing computer output, and recycling corrections.
Conversion Clerk:	Responsible for filing, controlling, retrieving, and matching sheets, keyed cards or sheets, and computer output. (for both Monographs and Serials)

SERIALS

Senior/Editor: Proofer	Responsible for supervising and coordinating editor-proofers.
2 Editors/Proofers:	Responsible for coding and editing sheets, proofing computer output, and recycling corrections.
(Conversion Clerk):	Duties shared by Conversion Clerk for Monographs.

2. SERVICES

Numbering:	Library codes and sequential numbers imprinted with numbering machine for refiling and linking records.
Reproduction:	Copies reproduced by Xerox machine of catalog cards, edited data, and inquiries to computer file.
Keying:	Tab cards keypunched (or scanner typed sheets) from conversion or maintenance data.
Computer Processing:	Original, interim, and post conversion data processed and secured on magnetic tapes.

3. ALLOCATION

Costs for systems development, and for conversion and maintenance of all files can be allocated to the Center, since each entry potentially can be held by any library.

Costs for computer preparation and printing of large Union catalogs once or twice during a decade can be allocated to the Center.

Costs for computer preparation and printing of frequent or local lists and catalogs can be proportionately allocated to libraries participating in the printing. One basis for such proportionate allocation can be calculated from computer tallies of: A = H/S, where

A = percentage of total printing costs allocated to each library participating in printing.

H = number of entries held by one library, out of the number of entries selected for print.

S = sum total of the number of entries held by all libraries and selected for print. (Note that this sum is not the number of unique entries printed, but of all the numbers held, without accounting for duplications.)

For example, suppose that three libraries participate in printing a selection of 100,000 entries from the master file, and that their respective holdings among those selected are: (a) 70,000; (b) 60,000; and (c) 80,000. Proportionate allocation of costs, then, would be:

(a) $\frac{70,000}{210,000} = 33\%$ (b) $\frac{60,000}{210,000} = 29\%$ (c) $\frac{80,000}{210,000} = 38\%$

Total printing costs are dependent on the number of characters actually expanded from the entries selected and printed, as enumerated in detail later in this volume.

4. SYSTEM DEVELOPMENT

The definition of this activity is simply the conversion of system design and program specification into a set of operating programs and procedures. In order to accomplish this conversion a three year effort is projected, consisting of three separate activities.

a. Programming: writing, debugging and running computer programs to accomplish the specified technical goals of the Center.

b. Training: developing instructional material for technical topics such as MARC, bibliographic programming, conversion editing/proofing, Processing Center file structure and organization.

c. Research: continuing analysis of topics which have high interest/pay-off for the Center, such as Automatic Field Recognition, subject-oriented data retrieval, acquisitions processing, network optimization.

The costs for this effort will be projected on a man year as well as a cost basis, to allow for variation in man year cost relative to different implementing agencies.

Since the effort is projected over three years, it is appropriate to establish annual performance benchmarks.

FIRST YEAR

Programming: Complete all input conversion programs, and format/print program for output. Complete substantial portion of MARC processing module. Start work on file maintenance. Start work on serials conversion and check-in modules.

Training: Develop courses for MARC conversion editing. Begin work on bibliographic programming material.

Research: Devote major effort to Automatic Field Recognition. Start work on subject access.

SECOND YEAR

Programming: Complete MARC processing and File Maintenance programs. Complete Retrieval programs. Begin work on Filing Rules and Authority Verification. Complete Serials conversion and check-in; begin work on Serials Binding, Claiming, and Union List programs.

Training: Complete Bibliographic Programming course. Begin Processing Center presentation for administrative personnel, regarding network philosophy and design.

Research: Complete Automatic Field Recognition. Continue Subject Access and Network Optimization.

THIRD YEAR

Programming: Complete Filing Rules and Authority Verification. Augment input and output programs to reflect research (Automatic Field Recognition) or new technology (photocomposition).

Training: Complete Network Courses

Research: Complete Subject Access, Network Optimization, Acquisitions studies.

(means begin task

) means end task

CHRONOLOGY OF SYSTEM DEVELOPMENT

YEAR	ACTIVITY	MONOGRAPHS	SERIALS
1st	Programming	(Input Conversion	(Input Conversion
		(Format/Print)	(Check-in
		(MARC Search	
		(File Maintenance	
	Training	(MARC Editing)	
		(Bibliographic Programming	
	Research	(Automatic Field Recognition	
		(Subject Access	
2nd	Programming	MARC Search)	Input Conversion)
		File Maintenance)	Check-in)
		(Retrieval)	(Binding
		(Filing Rules	(Claiming
		(Authority Verification	(Union List
	Training	Bibliographic Programming)	
		(Processing Center Design	
	Research	Automatic Field Recognition)	
		Subject Access)	
		(Network Optimization	
3rd	Programming	Filing Rules)	Binding)
		Authority Verification)	Claiming)
		(Augment input programs)	Union List)
		(Augment output programs)	
	Training	Processing Center Design)	
		(Processing Center Files)	
	Research	Network Optimization)	
		(Acquisitions)	

5. CONVERSION AND MAINTENANCE

Main phases of conversion and maintenance are listed in the following charts as abbreviated terms to delineate proposed types of time and effort for the initial three years of the Center. Such abbreviated terms cannot fully recapitulate the main terminology and many details of the whole report. Rather, the terms are meant to highlight the initial and major continuing tasks of the Center. Two main stages are designated: I. Initial Conversion, and II. Further Conversion and Maintenance.

A schedule has already been suggested for implementing Monographs and Serials during the initial 36 months of full production, in the earlier part of this Introduction under "Scope of Estimates" (see page 3). Accordingly, types of activity for Monographs and Serials will be described in the charts in parallel, rather than chronologically, to emphasize both the similarity and difference of the two implementations.

I. INITIAL CONVERSION

PHASE	MONOGRAPHS	SERIALS
A. Conversion Retrospective 1. Production & Control	Number all of catalog in one series for refiling, and main entries recto & verso in another series. Xerox copies with editing template except versos.	Number source records (probably Cardex, catalog card, shelf list, binding and accounting records). Xerox copies of each record.
2. Editing	Edit sheets and encode input information.	Select, transcribe, and encode input information on one sheet.
3. Keying	Key edited sheets in continuous string format (record number, card number, and data).	Key edited sheets in continuous string format.
4. Computer Processing	Enter & correct records, retain till approved, and then merge into Index and full data files.	Enter & correct records to establish Master and Local files.
5. Proofing	Verify formats and data, and recycle corrections.	Verify formats and data, and recycle corrections.
Interim (serving as parallel test and maintenance training)	Enter new records via MARC or full conversion (similar to retrospective)	Enter new records via MARC Serials (when available) or full conversion. Enter interim holdings via turn-about cards, verified directly from source records.

II. FURTHER CONVERSION AND MAINTENANCE

PHASE	MONOGRAPHS	SERIALS
B. Convert Catalogs of other libraries	Convert other catalogs via CSL-PC file or full retrospective and interim conversion.	Convert serials of other libraries via CSL-PC file or full retrospective & interim conversion.
C. Maintain established files	Enter new records via MARC or full conversion.	Enter new records via MARC Serials or full conversion. Maintain check-in of holdings via turn-about cards, and other serial functions as implemented for computer handling.
D. Authority Verification 1. Create authority files	Create and verify authority file with correct, complete, and standard author entries. Create and verify(with help of LC Subject Headings tape) authority file with correct, complete, and standard subject entries, and also "see" and "see also" references. Correct and resort Index and Bibliographical Master files, on basis of these standard author and subject entries.	Create and verify authority file with correct, complete, and standard titles-in-series with linkage for members of a common series. This will include all cross references of predecessors, parallel, & successors in time or accepted semantic variations, as well as any added entries. Correct and resort Master and Local files on basis of these cross references and few added entries.
2. Maintain authority files	Maintain relatively infrequent and small number of new entries and changes to author and subject authority files.	Maintain relatively few new serials and implied or separate changes in cross references or added entries.
E. Data Retrieval	Update author and subject entries before major printing. Consolidate local variations into Master record. Construct subfiles needed for printing.	Update cross references and entries before major printing. Consolidate local holdings & variations into Master record. Construct subfiles needed for printing.

6. OUTPUT SERVICES - CATALOGS

Book catalog production from machine readable bibliographic records involves five fundamental steps (starting with a fully updated file):

(1) Generate added entries and sort-keys, and resort the file.

(2) Convert the digital information on tape into graphic images, format each page, and produce a page image.

(3) Create offset master from the page image.

(4) Reproduce and collate multiple copies from page master on an offset press.

(5) Bind sheets into volumes, and prepare finished catalogs for distribution.

After describing these steps and some of the current capabilities available for doing them, we consider eight alternative cost options.

(1) EXPANSION AND SORT OF FILE

(a) Added entries:

In traditional card catalogs, added entries are continuously made so that the catalog will always be complete and ready for immediate retrieval. Any changes in author and subject headings entail manual revision of all related cards already in the file. In computer based files used for periodic printings, added entries need be made only at the times of printing, and all changes in author and subject headings are at the same time automatically brought up to date.

Generation of added entries is done entirely by the program, on the basis of explicitly specified conditions and the MARC codes in each record. These codes identify the data elements, and further help generate precedence and function codes for heading forms and types of entries. (Exceptions are manually coded, in the prior Authority Verification stage, for the small number of records which would otherwise conflict with the program because they are ambiguous, infrequent, or different.)

With these codes and program, as many or few added entries can be selected as needed, and the information later structured in the desired format.

(b) Sort-keys:

On catalog cards, sort-keys are the filing words typed in the fixed position above the author, except for the main entry. The type of entry is indicated by blank for the main entry, and by the case or color used in typing other kinds of entries. In a computer file, sort-keys also must be generated and identified in a fixed position within each record.

To make the sort-keys completely explicit, they are a combination of the desired data elements, with standardized editing of any words or characters to be disregarded or modified. Included in sort-keys are also precedence and function codes for heading forms and types of entry, plus explicit subordering needed for identical heading forms of the same type of entry. (See Vol.II FILE, pp. 96-171)

As with added entries, all sort-keys are generated completely by the software from MARC codes without any further manual effort.

(c) Sort expanded file:

With completely explicit sort-keys, and according to a specified collation sequence of characters, the main entries from the randomly ordered input file, as well as the newly inserted added entries, are sorted by a computer program into a file ready for conversion to page images.

(2) CONVERSION OF TAPE TO PAGE IMAGE

Three ways considered here to obtain a page image are:

(a) Computer Printout--printout from a computer line printer.

(b) CRT Microfilm--reduced size film from projection of a fixed character set on a CRT (cathode ray tube).

(c) CRT Photocomposition--page size film from construction of alternative character sets on a CRT.

The various aspects of producing a page image are described in terms of these three approaches.

(a) Computer Printout:

Digital to graphic conversion is most simply achieved by printing the tape on a computer line printer. Whatever is coded digitally on tape is automatically converted by hardware into the graphics available on a print chain. The effective rate for printing depends primarily on the number of distinct graphics on the chain used: upper case only, 60,000 lines/hr; both upper and lower case, 20,000 lines/hr. Diacritics and/or bold (weight) slow the print rate by a third, due to over printing of an estimated one third of all lines.

The formatting of each record depends on the format/print program. The line width per page is limited to 132 characters (including blanks within data, between columns, and allowing for margins). Vertical spacing (or leading) between lines is usually 6 lines/inch. Most printers can easily be adjusted to print 8 lines/inch, but the closeness of lines in the result decreases legibility. The data is usually lined up relative to the left margin. Proportional spacing is not possible since each character is of equal width (10/inch).

The output of the computer printing is a hard-copy, oversized page (11 x 14-7/8), ready for reduction in the next process to final page size desired.

(b) CRT Microfilming:

Digital to graphic conversion can also be done by CRT microfilming. In this method, the digital code on tape initiates (through software) a projection of an electron beam, through a stencil-like plate containing the character set (or graphics) onto a phosphorous coated facing. Each page image is built up on this facing, character by character, and then microfilmed.

Both upper and lower cases are available in the character set. Different point sizes of the same character are also available, made possible by the focusing electric and magnetic fields (controlled by software), which vary the intensity, enlargement ratio, and orientation of the projected characters. Special diacritics (or any character digitally coded on tape but not in the CRT character set) can be entered through software at a slightly increased cost.

(b) CRT Microfilm (Cont.)

Formatting of records on each page depends on software and pre-arranged codes. Line width per page is limited to 132 characters (as in computer printout), but 120 lines/page is possible (in contrast to 66 lines/page for computer printout, at 6 lines/inch), including headings and top and bottom margins. Proportional spacing is available, but justification of lines is not practical.

Output from the CRT is a 35mm negative film, with one page per frame. Efficient software will provide about 3600 frames/hr.

(c) CRT Photocomposition:

The third method to be described for producing a page image is CRT photocomposition. In this method, the tape is converted to one of the fonts already established for the CRT. Special characters, or diacritics, are also converted to "grid representations" (or graphics), and these graphics then linked to the codes on tape by a look-up table. Multiple fonts can be used as long as the magnetic tape has a unique code for each different graphic. (A bold upper case A, for example, and a regular upper case A must each have different codes on the tape. These unique codes, if not on the tape originally, can be introduced by software and an additional pass before going to the CRT.)

The formatting of the page again depends on the software and input codes. Photocomposition can create lines up to 11 inches (rather than 132 fixed characters), and justify most of these lines. The number of lines can be up to 120 per page.

The output is a page-size film, ready for transfer to a hard copy master.

(3) CREATION OF OFFSET MASTERS

For relatively small runs, offset masters are most economical. They can be created by electrostatic or photographic methods. In the electrostatic method (such as Xerox), the projected page image from the input copy is "fused" onto the master. In photography (photo-offset), the projected page image from the film negative is "burned" onto the master. The result of these two methods is the same as that achieved, without any reduction or enlargement, by typing directly on a master. The photographic process allows two frames (page images) to be burned onto an 11" x 17" master. Even though the larger sheet must be folded (or cut), reproduction costs are greatly reduced.

The offset master is usually a heavy stock, hard finish paper that will yield up to 100 reproductions from the electrostatic method, and up to 500 reproductions of better quality from photography. For greater quantity, additional paper masters can be readily created.

The method chosen for creating the master depends on the final page-size and quality desired, and on how the original page image was produced:
(a) From computer printout the offset master can be created by reducing the oversized page to the desired size (by photography), or to 8 1/2 x 11 size (by Xerox Computer Forms Processor).
(b) From CRT microfilm the offset master can be created by enlarging the 35mm microfilm to the desired size (by photography), or to either 8 1/2 x 11 or 11 x 14 7/8 size (by Xerox Copy Flow 11 Continuous Printer).
(c) From CRT photocomposition the page size film is transferred photographically onto the desired size of master.

(4) REPRODUCTION AND COLLATION OF SHEETS

Once the offset masters are created, reproduction of multiple copies (on quality of paper desired) can be done on any offset press. Cutting and collation of the reproduced sheets then follow, with some presses collating automatically, at least for a limited number of copies.

(5) BINDING AND DISTRIBUTION OF VOLUMES

Final preparation includes design and printing of the covers, binding of individual volumes (with a recommended maximum of 300 sheets, or 600 pages per volume, for durability), packaging, and distribution of completed catalogs.

ALTERNATIVE COST OPTIONS

To illustrate the spectrum of options, eight alternatives are described in terms of the different methods within five steps. For cost comparison, these steps are assumed to be done as:

Step 1. (generating entries and sorting file), by CSL's Computer Center.
Step 2 and 3 (converting tape to page image, and creating an offset master) involve the options which distinguish the eight alternatives, as outlined in the following chart:

Abbreviated Alternatives	*Distinctive Machine	Cases used in font	Step 2 Magnetic Tape to Page Image	Step 3 Page Image(s) on Offset Master
CPL =	Computer Line Printer	Lower/ Upper	Computer Printout	Photographic
CPU =	Computer Line Printer	Upper only	Computer Printout	Photographic
CXL =	Computer Line Printer	Lower/ Upper	Computer Printout	Electrostatic (1)
CXU =	Computer Line printer	Upper only	Computer Printout	Electrostatic (1)
SC1 =	Stromberg Carlson 4060	Full Font	CRT- Microfilm	Photographic
SC2 =	Stromberg Carlson 4060	Full Font	CRT- Microfilm	Electrostatic (2)
SC3 =	Stromberg Carlson 4060	Full Font	CRT- Microfilm	Photographic
SV =	Sedgwick Videocomp	Multi Font	CRT- Photo composition	Photographic

*Distinctive Machine used for method in Step 2.

(1) Xerox Computer Forms Processer
(2) Xerox Copy Flow 11 Continuous Printer.

Step 4 (reproducing and collating sheets), by outside duplicating service, except in option (SC1)--George's Litho--which carries the entire printing process from sorted magnetic tape to collated catalog sheets, binding, and preparing volumes for distribution.

Step 5 (binding and distributing catalogs), shared by Binder and CSL distribution, or done entirely by Binder with list furnished by CSL.

Cost Summaries, in terms of these eight alternatives, are presented in the next section of this report, while the basis for such costs are contained in Appendix IV.

C O S T S U M M A R I E S

A. SYSTEMS DEVELOPMENT

B. ESTABLISHMENT OF CENTER, CONVERSION AND MAINTENANCE

C. UNIT COST (CONVERSION AND MAINTENANCE)

D. OUTPUT SERVICES

A. SYSTEM DEVELOPMENT

A summary of the projected development budget is given in both man years and dollars. The dollar figures are given at the scale of:

$30,000 per Senior Technical Man Year
$25,000 per Intermediate Man Year
$20,000 per Junior Technical Man Year

These are hypothetical costs comparable to those charged by software houses.

Other figures may be used, if other rates are available. Machine Time costs were estimated at $100 per hour.

A more detailed, year by year, task by task breakdown of this budget will be found in Appendix II - System Development.

SUMMARY OF DEVELOPMENT BUDGET

Task	Year 1	2	3	Total Man Years		Total Dollars
Programming	6	4.5	3.5	14	@25K	$350,000
Training	1.5	1.5	.5	3.5	@20K	70,000
Research	2	2	1.5	5.5	@25K	137,500
Management	2	2	2	6	@30K	180,000
Total Man Years	11.5	10	7.5	29		
Total Man Cost	$290,000	$252,500	$195,000			$737,500
Machine Cost	$ 21,000	$ 14,400	$ 12,000			$ 47,400
Grand Total	$311,000	$266,900	$207,000			$784,900

B. ESTABLISHMENT OF CENTER, CONVERSION, AND MAINTENANCE

A summary for this projected part of the budget is given as total costs allocated to:

(1) Personnel, permanent and additional
(2) Services and Expenses
(3) Recap of Personnel, Services, and Expenses

Estimated costs rates and amounts of material are stated where applicable.

A detailed, task-by-task breakdown of this summary will be found in Appendix III, Conversion and Maintenance.

ESTIMATED IMPLEMENTATION COSTS
TO ESTABLISH CSL PROCESSING CENTER AND TO CONVERT AND MAINTAIN FILES
(Six months preparatory; three years full operation)

PERMANENT PERSONNEL:	Prep-	Yr. 1	Yr. 2	Yr. 3	Total
Center Coordinator @ 15,000/yr.	$ 7,500	$ 15,000	$ 15,000	$ 15,000	$ 52,500
Operations Manager @ 12,000/yr.	6,000	12,000	12,000	12,000	42,000
Data Supervisor @ 10,000/yr.	2,500	10,000	10,000	10,000	32,500
Technical Editor @ 7,500/yr.	0	7,500	7,500	7,500	22,500
Secretary @ 7,000/yr.	3,000	7,000	7,000	7,000	24,000
2 Production Clerks @ 6,500/yr.	0	13,000	13,000	13,000	39,000
5% Increases + 10% Benefits	1,900	6,450	10,000	13,720	32,070
Total Permanent Personnel	20,900	70,950	74,500	78,220	244,570
ADDITIONAL PERSONNEL (MONOGRAPHS)					
2 Senior Editor/Proofers @ 9,000/yr.	4,500	18,000	18,000	18,000	58,500
8 Conversion Editor/Proofers @ 8,000/yr.	0	64,000	64,000	64,000	192,000
Conversion Clerk @ 6,000/yr.	0	6,000	6,000	6,000	18,000
5% Increases + 10% Benefits	450	8,800	13,640	18,720	41,610
Total Addtl. Pers. (Monogr.)	4,950	96,800	101,640	106,720	310,110
ADDITIONAL PERSONNEL (SERIALS)					
Senior Editor/Proofer @ 9,000/yr. (2-1/4 yrs.)	0	2,250	9,000	9,000	20,250
2 Conversion Editor/Proofers @ 8,000/yr. (1 in 2nd yr., 2 in 3rd yr.)	0	0	8,000	16,000	24,000
5% Increases + 10% Benefits	0	225	2,635	5,320	8,180
Total Addtl. Pers. (Serials)	0	2,475	19,635	30,320	52,430
Total Initial Salaries	23,500	154,750	169,500	177,500	525,250
Total Increases (5%/yr.)	0	0	8,475	18,190	26,665
Total Employee Benefits (10%)	2,350	15,475	17,800	19,570	55,195
TOTAL DIRECT PERSONNEL COSTS	25,850	170,225	195,775	215,260	607,110

CONVERSION AND MAINTENANCE

SERVICES AND EXPENSES

SERVICES:	Yr. 1	Yr. 2	Yr. 3	TOTAL
Numbering Machine Lease @ $1,000/yr.	$ 1,000	$ 1,000	$ 1,000	$ 3,000
Xeroxing @ $0.025/sheet				
Full conversion	5,940	5,280	4,950	16,170
Partial conversion (1.32 sheets/record)	0	2,475	4,950	7,425
Total Xeroxing	5,940	7,755	9,900	23,595
Keying @ $0.0715/tab card				
Full conversion (5/record)	64,350	57,200	53,625	175,175
Partial conversion (2/record)	0	10,725	21,450	32,175
Corrections	9,220	7,220	6,990	23,430
Authority verification	5,360	1,970	1,140	8,470
Total Keying	78,930	77,115	83,205	239,250
Computer Processing @ $100/hr.				
Initial conversion	29,700	14,000	0	43,700
Other conversions	14,600	29,200	43,800	87,600
Maintenance	6,000	13,000	21,000	40,000
Authority verification	2,550	2,190	1,940	6,680
Data retrieval	3,580	7,800	11,780	23,160
Total Computer Processing	56,430	66,190	78,520	201,140
TOTAL SERVICES	142,300	152,060	172,625	466,985
EXPENSES				
Transportation: weekly trips to Computer Center @ $5/trip	130	260	260	650
Field trips, 2 people, 25 trips/yr. @ $40/trip	1,000	1,000	1,000	3,000
Data Base Supplies:				
Xerox paper @ $1.20/ream 1890 reams	575	745	950	2,270
Computer tape reels @ $20/reel, 10 reels/month	2,400	2,400	2,400	7,200
Documentation @ $3/page	2,100	2,100	2,100	6,300
External 400 pages/yr.				
Internal 300 pages/yr.				
Space preparation (Price not estimated)		-	-	-
TOTAL EXPENSES				
Preparatory $1,000	6,205	6,505	6,710	20,420
TOTAL SERVICES AND EXPENSES				
Preparatory $1,000	148,505	158,565	179,335	487,405

CONVERSION AND MAINTENANCE

RECAP OF PERSONNEL, SERVICES, AND EXPENSES

	Pre-	Yr. 1	Yr. 2	Yr. 3	Total
Personnel	$25,850	$170,225	$195,775	$215,260	$ 607,110
Services	0	142,300	152,060	172,625	466,985
Expenses	1,000	6,205	6,505	6,710	20,420
Total Direct Costs:	26,850	318,730	354,340	394,595	1,094,515
Total Indirect Costs: (20% Personnel for overhead & other library operations)	5,170	34,045	39,155	43,055	121,425
TOTAL COSTS	$32,020	$352,775	$393,495	$437,650	$1,215,940

COMPARISON OF DERIVATION OF UNIT COSTS:

Applicable direct costs of converting monographs, before benefits:

	Pre-	Yr. 1	Yr. 2	Yr. 3	Total
Personnel:					
Production Clerks	$ 0	$ 13,000	$ 13,000	$ 13,000	$ 39,000
Senior Editors	4,500	18,000	18,000	18,000	58,500
Regular Editors	0	64,000	64,000	64,000	192,000
Conversion Clerk		6,000	6,000	6,000	18,000
Total Personnel:	4,500	101,000	101,000	101,000	307,500
Services:		142,300	152,060	172,625	466,985
Expenses:		6,205	6,505	6,710	20,420
Total Monograph Cost:	$ 4,500	$249,505	$259,565	$280,335	$ 794,905
Derived Personnel + Services + Expenses Unit Cost:		$ 1.411	$ 1.105	$ 0.934	
compared to:					
Operationally derived Unit Cost: (see derivation next page)		$ 1.343	$ 1.058	$ 0.925	

First derivation is larger in first two years, due to allowance made for extra volume of original conversion and possible increase in number of corrections, above stated volumes of 180,000; 235,000; 300,000; and 2% errors.

C. UNIT COST SUMMARY

A summary is given of unit costs derived in converting and maintaining monographs during full operation after the initial six-month period of preparation.

Costs are based on the basic five phases of:

A. Initial Conversion.
B. Conversion of other libraries.
C. Maintenance of files (MARC)
D. Authority verification.
E. Data retrieval.

UNIT COSTS -- FULL OPERATION
(CONVERSION AND MAINTENANCE)

DESCRIPTION	UNIT COST (rounded cents)	YEARS 1 Volume	1 Total Cost	2 Volume	2 Total Cost	3 Volume	3 Total Cost
Initial Conversion							
*MONOGRAPHS							
A-1. Production & Control	$ 0.12	180,000	$ 21,600	85,000	$ 10,200	0	$ 0
A-2. Editing	0.26	180,000	46,800	85,000	22,100	0	0
A-3. Keying	0.36	180,000	64,800	85,000	30,600	0	0
A-4. Computer Processing	0.18	180,000	32,400	85,000	15,300	0	0
A-5. Proofing	0.26	180,000	46,800	85,000	22,100	0	0
Total Monographs Initial Conversion	1.18	180,000	$212,400	85,000	$100,300	0	0
Conversion of other libraries (Duplicate holdings)							
B-1. Production & Control	$ 0.12	0	0	75,000	$ 9,000	150,000	$ 18,000
B-2. Editing	0.09	0	0	75,000	6,750	150,000	13,500
B-3. Keying	0.12	0	0	75,000	9,000	150,000	18,000
B-4. Computer Processing	0.03	0	0	75,000	2,250	150,000	4,500
B-5. Proofing	0.04	0	0	75,000	3,000	150,000	6,000
Total conversion other libraries (dupl. monographs)	$ 0.40	0	0	75,000	$ 30,000	150,000	$ 60,000
Total conversion other libraries (unique monographs)	$ 1.18	0	0	75,000	$ 88,500	150,000	$177,000
C. Maintenance of files (MARC)	$ 0.10	60,000	$ 6,000	130,000	$ 13,000	210,000	$ 21,000
D. Authority Verification	Yr. 1 0.11	180,000	19,800	-	-	-	-
	2/3 0.07	-	-	122,500	8,575	75,000	5,250
E. Data Retrieval	0.02	180,000	3,600	415,000	8,300	715,000	14,300
		Added Volume	Total Cost	Added Volume	Total Cost	Added Volume	Total Cost
TOTALS		180,000	$241,800	235,000	$248,675	300,000	$277,550
Average Yearly Unit Cost		$1.343		$1.058		$0.925	

*SERIALS on next page

SERIALS UNIT COSTS -- FULL OPERATION (CONVERSION AND MAINTENANCE)

NOTE: Conversion costs are additive, i.e. apply only to additional serials converted.
*Maintenance costs are cumulative, i.e. apply to all serials converted to date.
Maintenance includes costs within CSL Center, not within each local library.

DESCRIPTION	UNIT COST (rounded cents)	YEARS 1 Volume	1 Total Cost	2 Volume	2 Total Cost	3 Volume	3 Total Cost
A. Initial Conversion Estimated on proportionate basis of monographic effort.	4:1 in data; 1.1:1 in processing.	1st library Initial group of serials for test & acceptance of ongoing one library system.		1st library Expanded group of serials for test&acceptance of ongoing two library system.		1st library Complete group of serials for test&acceptance of ongoing many library system.	
Total Serials Initial Conversion	$5.19	500	$ 2,595	3,500	$18,165	2,000	$10,380
B. Conversion of other libraries Incorporating advances in serials at national level and in local systems development.		-	-	2nd library for test & acceptance of two schedules and sets of local data.		2nd & other libraries for test&acceptance of many schedules & local data.	
Dups(title in system) plus local data	$1.32	-	-	600	$ 790	9,400	$12,410
New titles to system plus local data	$5.19	-	-	300	$ 1,555	1,700	$ 8,825
*C. Maintenance (Cumulative) 1st year	$1.09	500	$ 545	-	-	-	-
2nd year	$0.49	-	-	4,900	$ 2,400	-	-
3rd year	$0.44	-	-	-	-	18,000	$ 7,920
D. Authority Verification New	$0.48	500	$ 240	4,100	$ 1,970	3,700	$ 1,775
Dup	$0.31	-	-	300	95	9,400	2,915
E. Data Retrieval	$0.09	500	$ 45	4,400	$ 395	13,100	$ 1,180
Added Volume: 8,000 unique 10,000 duplicate		(500)		(4,400)		(13,100)	
TOTAL VOL. & COSTS		500	$ 3,425	4,900	$25,370	18,000	$45,405
Average Yearly Unit Cost of cumulative volume		$6.850		$5.178		$2.523	

D. OUTPUT SERVICES

Summary of costs for producing book catalogs by eight different methods from original computer tape file is given in terms of the five basic steps to:

(1) Add entries and sort file.
(2) Convert digital codes on computer tape to graphics, the equivalent of a typeset page.
(3) Create offset master of these page images.
(4) Reproduce, cut and collate desired number of copies from master, ready for binding.
(5) Bind sheets and prepare catalogs for distribution.

Specific companies have been consulted for cost estimates, but it is intended that their estimates reflect only comparative rates for the eight methods. (For further information relating to these methods, as well as details pertaining to the cost analysis, see Appendix IV.)

Cost of the first step (adding entries and sorting file) is common to all methods. Cost of steps 2 and 3 (first page creation) varies according to the method chosen, but reproduction and binding costs (steps 4 and 5) are independent of method and vary only with the number of copies produced.

The costs of eight alternative methods for the five steps are presented in both tabular and graph form. Finally, estimated unit costs for 100,000 titles are displayed.

Explanation of Factors Involved in the Cost Analysis

A fixed page format (involving margins, column gap lengths, and vertical and horizontal print lengths) was assumed and together with a specified print density these factors determined the number of characters/page and the number of pages/copy. In considering the average number of characters in the different entries, we decided that the bibliographic record for author main entry, subject entries, and any series entry would appear in full form, while the bibliographic record for author added entries and title entries would be displayed only in part (an average of 0.6 x full entry).

Specific values affecting and assumed in this analysis are listed below:

Avg. number of characters/title, input (MARC II)	580
Avg. number of characters/main entry, output	225
Avg. number of subject entries/title	1.4
Avg. number of author added entries/title	0.2
Avg. number of title entries/title	1.0
Avg. number of series entries/title	0.05
Avg. number of characters/title, output (main and added entries)	741
Computer cost/hr	$100

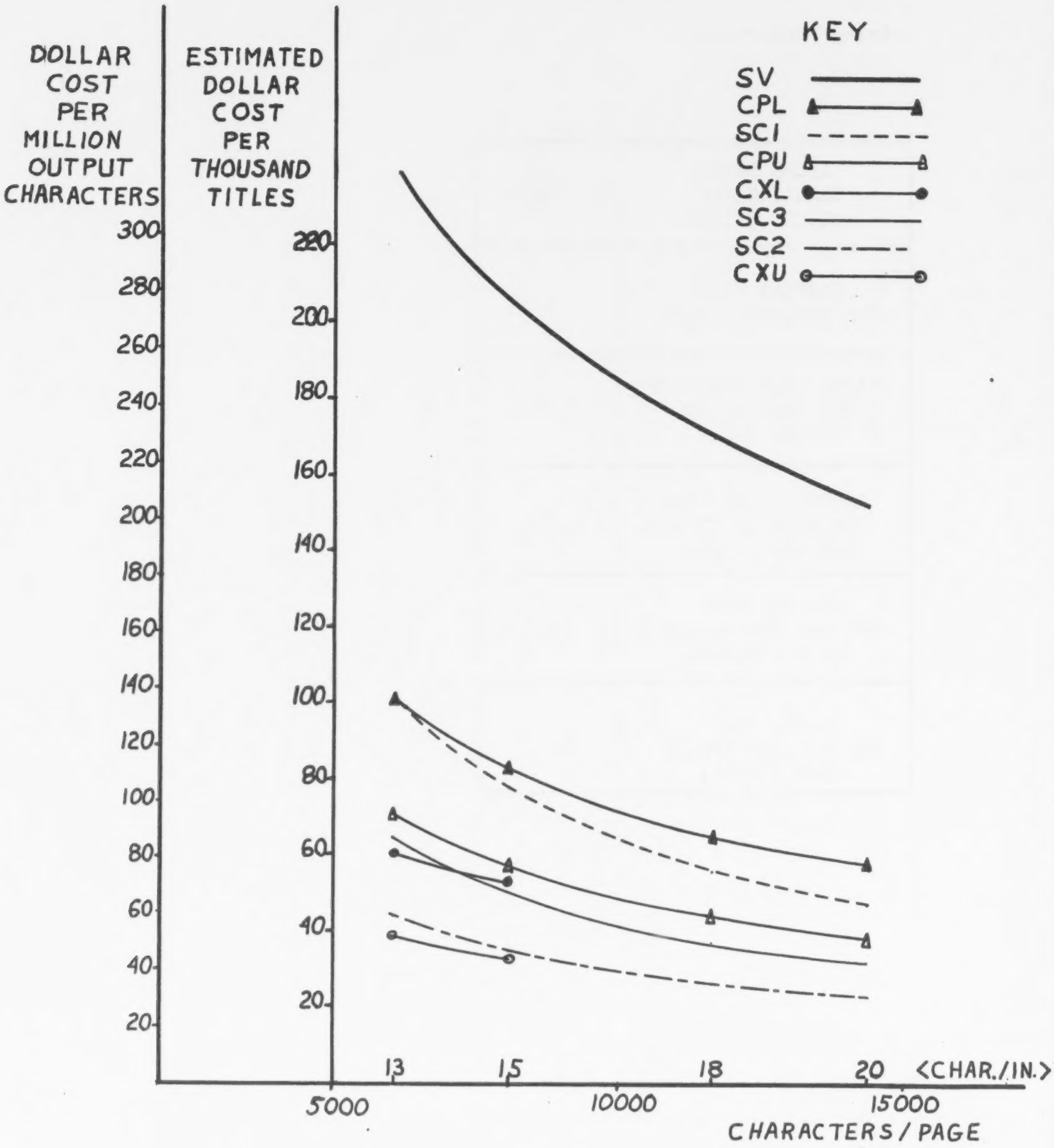
The comparative analysis for CXL and CXU is only valid for up to approximately 8000 characters/page since that amount represents the maximum amount of characters obtainable on a 11" x 14 7/8" computer form. For CPL and CPU the proper ratio of characters per line versus lines per reduced page was determined so as to create the densities (after photo reduction) corresponding to the other methods.

STEP			SC1	SC2	SC3	SV	CXU	CXL	CPU	CPL
1	ADDED ENTRY CREATION PER 1000 TITLES		2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
	SORTING FILE (1) PER 1000 TITLES		4.91	4.91	4.91	4.91	4.91	4.91	4.91	4.91
2 and 3	FIRST PAGE CREATION COST PER 1000 TITLES	13 C/in 7.8 L/in	93.92	36.59	57.33	231.20	31.45	52.91	63.78	93.90
		15 C/in 9.0 L/in	70.55	27.49	43.06	197.41	25.35	44.93	49.70	75.81
		18 C/in 10.8 L/in	48.99	19.09	29.90	161.87	---	---	36.27	58.01
		20 C/in 12 L/in	39.68	15.46	24.22	144.53	---	---	30.59	50.49

STEP			$\frac{13}{7.8}$	$\frac{15}{9.0}$	$\frac{18}{10.8}$	$\frac{20}{12.0}$
	$\left\{\frac{\text{Characters/inch}}{\text{Lines/inch}}\right\}$					
4	REPRODUCTION COST PER MILLION CHAR. PER COPY		0.470	0.353	0.245	0.198
	REPRODUCTION COST PER THOUSAND[2] TITLES PER COPY		0.348	0.261	0.181	0.147
5	BINDING COST PER 1000[3] TITLES	100 Copies	14.00	10.55	7.34	5.97
		200 Copies	21.70	16.35	11.37	9.25
		300 Copies	27.70	20.90	14.51	11.80
		400 Copies	32.00	24.20	16.80	13.67
		500 Copies	36.70	27.70	19.25	15.65
	BOOK CATALOG[4] PAGES PER 1000 TITLES		122	92	64	52

1. Used sorting rate for 100,000 titles.
2. Assuming 2 page images on 11" x 17" master.
3. Rates listed are for 300 sheets/volume (600 pages), cover printing, packaging.
4. Assuming $8\frac{1}{2}$ x 11 inch paper with two columns.

ADDED ENTRY GENERATION, SORTING, AND FIRST PAGE CREATION COST



Unit Cost Estimate

For 500 copies of 100,000 titles at print density 18 characters/inch (and 10.8 lines/inch) an upper limit unit cost estimate is given for book catalog production.

ADDED ENTRY CREATION PER 100,000 TITLES	$ 208.
SORTING FILE PER 100,000 TITLES	491.
FIRST PAGE CREATION COST PER 100,000 TITLES (SC1)	4,899.
REPRODUCTION COST PER 100,000 TITLES PER 500 COPIES	9,070.
BINDING COST PER 100,000 TITLES PER 500 COPIES	1,925.
TOTAL COST PER 100,000 TITLES (500 COPIES)	16,593

A P P E N D I C E S

I. Reprint of PRELIMINARY ORGANIZATIONAL DESIGN (with minor revisions)

II. SYSTEMS DEVELOPMENT

III. CONVERSION AND MAINTENANCE

IV. OUTPUT SERVICES

Appendix I

PROCESSING CENTER ORGANIZATIONAL DESIGN

(Vol. II California State Library: Processing Center Design and Specifications, pp. 253-258; minor revisions made in this reprint)

I. PRELIMINARY ORGANIZATIONAL DESIGN: INTRODUCTION

This section covers preliminary specifications for setting up and developing full-scale operation of the computer based California State Library Processing Center (CSL-PC). Specifications which follow are grouped under two headings: A) The operations of the center and its organization; B) The implementation stages of the center with a proposed overall schedule for carrying out its work.

* * *

A. OPERATIONS AND ORGANIZATION OF CSL-PC

Organizational design is considered here as having two general aspects: the overall plan of operations to be carried out, and the staffing, equipment, and facilities needed to carry out such a plan. Each aspect is discussed separately below.

OVERALL PLAN OF OPERATIONS. The plan proposed here for CSL-PC is aimed at full scale implementation in three basic areas of library functions: Cataloging, Serials, and Acquisitions, plus additional functions developed later. Each area, of course, will involve various phases. Such a plan is intended to reflect a workable interpretation of the policy and goals decided upon by California State Library, from initial to full scale operation.

Full scale operation can be considered in either the immediate or long range sense. The immediate is the completion of all the phases within the first area of Cataloging, defined as full conversion and maintenance for producing the book catalog. Full scale operation in the long range is the completion of all three areas in all their phases; for example, greater usage of MARC tapes in technical processing.

The immediate sense of full scale operation for Cataloging is the main emphasis here. Serials can be started later and developed independently, even if concurrent with some of the programming in the first area. Acquisitions can be developed dependent upon the success in Cataloging.

The actual order of implementation leading to full-scale operation will depend first on when each area is to be initiated, and second on the effective scheduling of phases once an area is started.

STAFFING, EQUIPMENT, AND FACILITIES. Staffing, equipment, and facilities for CSL-PC (and outside services) will reflect what stage of operation has actually been reached. The organizational design of CSL-PC in its internal workings is proposed as relatively stable, at least for the first area of Cataloging. It will expand somewhat as more areas are implemented.

In the conversion stage of Cataloging, the administrative personnel will be doing the same types of activity as in maintenance, i.e. designing and directing plans of implementation. Operational personnel in the conversion stage will be mostly preparing source copies of the huge volume of catalog cards for outside editors, and assisting in the final control of converted data. In the maintenance stage, the same operational personnel will switch over and absorb the functions of editing and proofing a much smaller volume of changes going to and from outside keying and computer processing in Sacramento.

The organizational design for outside services in the conversion stage is for editing and proofing, keying, and initial programming and conversion computer processing to be done at the place where personnel are available and the work can be done most accurately and economically.

In the maintenance stage, the keying can continue to be done by an outside service (until such time as there is enough computer keying to warrant a full time keying operator and machine). The keying can be done by the same firm that does the computer processing or by a keying service within easy transportation of both the CSL-PC and the computer processing center.

STAFFING: ADMINISTRATIVE. Administrative persons are considered those who are involved in directing and planning implementation of overall policy and goals.

Center Coordinator, head of the center: He or she should be a person knowledgeable about systems analysis and the technical aspects of data processing, and also about the procedures and bibliographical data of libraries.

His chief responsibility is carrying out decided Cal State policy and goals on data processing, and reporting back on its implementation. His chief duty is in developing and coordinating the plan of implementation as carried out by the center and outside services, and other library operations.

Secretary to the Center Coordinator: Responsible for secretarial work and documentation done by the Coordinator, and other secretarial work of the center.

STAFFING: OPERATIONAL. Operational persons are considered those who are involved in supervising and carrying out actual operations within the center, or operations interconnected with other library operations or outside services.

Operations Manager, supervisor of the center: Operational responsibilities for work and documentation done by the center, and within the library's manual work as it relates to data processing.

Data Supervisor: A librarian with cataloging experience should be the Supervisor for conversion of both monographs and serials. In addition to overall supervision, this person would resolve editing and proofing questions, and be responsible for training new editors in the use of coding manuals and proofing procedures.

Technical Editor for the Operations Manager: Responsible for reports and documentation done by operational personnel.

Clerks during conversion: Responsible for selecting, numbering, reproducing, and filing source copies; and assisting in interim and final control/recovery of source data.

Clerks after conversion: Responsible, as before, for the preparation of source copies, as well as absorbing the editing, proofing, and control of all documents. With a reduced volume of cards, less pressure in time, and possible upgrading and training, the additional editing and proofing can be handled without highly trained personnel.

STAFFING: OUTSIDE PERSONNEL.

Programmers from software development computer facility (State agency or otherwise): They are responsible for coding, testing, and documenting programs based on the original design specifications. Total number of programmers depends very much on experience, background and organization of the software development staff.

A single programmer can start the work, and then be joined by another as the work accelerates or a programmer becomes available. A single programmer, however, for the total years of effort necessary for complete development would not only delay final production, but would also be risky since the programmer could leave. An enormous amount of unresolved and unrecorded analysis is known only to the programmer until the program is tested and documented. Any abrupt change in programmers can be a substantial loss in effort and time.

Senior Editor/Proofers: These should be persons responsible for preparatory and continuing editing and proofing, immediate supervision of editor/proofers, coordination of receipt of coding sheets to and from keying and computer processing, and channeling of editorial problems to Data Supervisor.

Editors/Proofers: These should be persons capable of doing detail editing and proofing work, and, preferably, with some knowledge of cataloging. Our experience indicates that library science students are suitable. The number of editors will depend on the projected schedule for completing the conversion.

Conversion Keyers: The number of keyers would be determined by the volume of coding sheets turned out by the editors. The keyers would key all data according to prepared instructions of required format

and special symbols, and return coding sheets and keyed media to editors/proofers.

Computer Operators/Data Clerks: Obtained from the computing center. Responsible for processing and securing conversion and maintenance data.

EQUIPMENT. Equipment noted here is special to the center and its operations, and does not include regular equipment normal to most offices, such as typewriters, cabinets, furniture, etc. It also does not include a computer, which is not owned or operated by the center during initial operation.

Numbering machine: A Cummins 270 series machine or Pitney Bowes machine, with continuous numbering can be leased at about $1,400 per year, or purchased at slightly more than three years' lease.

Reproduction machine: A Zerox 720 machine (lease only) costs 2 1/2¢ a copy, plus materials and labor. (The Xerox 3600 with hopper has also just come on the market.) This cost figures a minimum number of copies per month, a number well below the anticipated production of the Processing Center.

Two work tables and one storage cabinet: Required during conversion. One table for numbering and selecting, the other for assembling completed copies which are stored in the cabinet in flat reams of paper.

One tab card and one document cabinet: Required during maintenance for storing keyed data and computer printouts for verification and reference.

FACILITIES: SPACE FOR PERSONNEL. Basic space will be required for the administrative head and secretary, for operations manager and typist, and for two operational clerks working at desks during conversion on documentation and control, and in addition on editing and proofing during maintenance.

FACILITIES: SPACE FOR EQUIPMENT. Space will be needed for work table and numbering machine, preferably near the catalog during numbering, selecting, and refiling cards.

Space of relatively noise-contained room will be needed for reproduction machine, along with the assembly table and storage cabinet for completed packages of source copies.

B. STAGES OF IMPLEMENTATION

Stages of implementation are discussed here from the point of view of the supervision of the total process in which the stages actually occur. The effectiveness of these actual stages is considered dependent upon initial scheduling and rescheduling where appropriate.

INITIAL SCHEDULING. Implementation can begin in Cataloging, with Serials implemented independently. In order to have direct access to MARC II tapes, Acquisitions should be started after the cataloging phase is completed.

The start and termination of any implementation, of course, is dependent on the manpower and funds available, on the successful completion of logically prior phases, and on the inherent complexity of procedures involved and their potential interrelation in a unified dynamic system. Once the goals are set, the rate and the order in which a stage should be implemented will depend on resources, actual performance, and the complexities of procedures being implemented.

It is recommended that implementation in any phase or area wait until separate programming and supporting manual work are feasible and can be done without delaying implementation of any phase already started.

RESCHEDULING. A schedule for implementation will include periodic review, report, and adjustment as the work proceeds. Such reviews serve to combine the functions of monitoring, documenting, and (where necessary) re-estimating and rescheduling parts of each job.

APPENDIX II: SYSTEM DEVELOPMENT COSTS

Summary

This appendix presents separate task/phase cost projections for the CSL-PC System Development effort. The breakdown is by time across these separate functions:

A. Programming

B. Training

C. Research

D. Management

Within each separate analysis, an attempt is made to project subtasks and functions.

A. Programming (cost @ $25,000 per man year)

Function	Year 1	2	3	Total Man Yrs.	Total Cost
Input	1.5		.5	2	50,000
Print/Format	1		.5	1.5	37,500
MARC	1	.5		1.5	37,500
File Maint	.5	1		1.5	37,500
Retrieval		.5		.5	12,500
Filing Rules		1	.5	1.5	37,500
Authority Work		1	.5	1.5	37,500
Serials	.5	1.5	2	4	100,000
Total m.y.	4.5	5.5	4	14	
Total $	112,500	137,500	100,000		350,000

B. Training (cost @ $20,000 per man year)

Function	Year 1	2	3	Total Man Yrs.	Total Cost
MARC editing	1			1	20,000
Bibliographic Programming	.5	.5		1	20,000
Processing Center	___	1	.5	1.5	30,000
Total m.y.	1.5	1.5	.5	3.5	
Total $	$30,000	$30,000	$10,000	$70,000	$70,000

C. Research (cost @ $25,000 per man year)

Function	1	2	3	Total Man Yrs.	Total Cost
A.F.R.	1.5	.5		2	50,000
Subject Access	.5	1		1.5	37,500
Network		.5	.5	1	25,000
Acquisitions			1	1	25,000
Total M.Y.	2	2	1.5	5.5	____
Total $	$50,000	$50,000	$37,500	$137,500	$137,500

D. Management (cost @$30,000 per man year)

Function	1	2	3	Total M.Y.	Total Cost
Programming	1	1	1	3	90,000
Research	.5	.5	1	2	60,000
Training	.5	.5		1	30,000
Total M.Y.	2	2	2	6	
Total $	$60,000	$60,000	$60,000	$180,000	$180,000

Conversion and Maintenance

APPENDIX III: ESTABLISHMENT OF CENTER, MAINTENANCE, AND OTHER CONVERSIONS

SUMMARY:

This appendix presents separate task/phase cost projections for:

A. Initial Conversion:

MONOGRAPHS

1. Production and Control
2. Editing
3. Keying
4. Computer Processing
5. Proofing

SERIALS

6. Complete Initial Conversion (stated proportionately)

B. Conversion of other libraries

C. Maintenance of established files

D. Authority Verification

E. Data Retrieval

SUMMARY: MONOGRAPHS--PRODUCTION AND CONTROL (Initial Conversion)

This Section A-1 presents estimated costs to:

Number all cards on verso for refiling.

Select, and number main entries, recto and verso.

Reproduce main entries, recto.

Select, count, and reproduce versos of main entries.

Interfile, and package reproduced copies of main entries, verso and recto.

Refile, and recount main entry cards.

Costs are based on:

(1) Retrospective conversion.

(2) Interim changes during conversion.

(3) Basic formulas and values assumed in these costs.

ESTIMATED TIME AND COST FOR PRODUCTION AND CONTROL

Phase A1: Retrospective (Initial Conversion)

See page 60 for the formulas used here -

JOB	STEP	VOLUME/UNIT	TIME Rate per/hr	COST Rate per/hr	1.25 TOTAL TIME	1.25 TOTAL COST
1. NUMBER all catalog cards	Set up/ Number	500 runs 500,000 cards	5 runs 15,000	$5 $3	125 hrs 42	$ 625 $ 126
*2. SELECT main entries	Select/ Verify	500,000 cards 250,000 cards	500 500	$3 $4	1250 625	$ 3,750 $ 2,500
3. TRANSPORT cards		250 runs	5 runs	$3	63	$ 189
4. NUMBER main entries, front & back	Set up/ Number	500 runs 500,000 cards	5 runs 15,000	$5 $3	125 42	$ 625 $ 126
5. REPRODUCE main entry fronts		250,000 cards	200	$8.48	1563	$13,255
6. SELECT verso cards	Select/ Verify	250,000 cards 195,000 cards	500 2000	$3 $3	625 122	$ 1,875 $ 366
7. COUNT verso cards	Set up/ Count	250 runs 55,000 cards	5 runs 15,000	$5 $3	63 5	$ 315 $ 15
8. REPRODUCE verso cards		55,000 cards	200	$8.48	344	$ 2,918
9. INTERFILE verso sheets		55,000 sheets	700	$3	99	$ 297
10. PACKAGE & label sheets		305,000 sheets	20 pkg.= 10,000 sheets	$3	39	$ 117
11. REFILE main entry cards		250,000 cards	700	$3	447	$ 1,341
12. COUNT refiled cards	Set up/ Count	250 runs 500,000 cards	5 runs 15,000	$5 $3	63 42	$ 315 $ 126
				RETROSPECTIVE TOTAL:	5684 hrs	$28,881

* Serial main entries also can be sub-selected at this point. (See Volumes and Rates)

Assume two full time clerks for this procedure for 18 months; thereafter, as upgraded assistants for full file maintenance, and further conversion.

Crucial verification of selection of main entries can be delegated to a librarian acquainted with CSL catalog, or partially absorbed by Operations Manager.

ESTIMATED TIME AND COST FOR PRODUCTION AND CONTROL:

Phase A1: Interim (Initial Conversion)

JOB	STEP	VOLUME/UNIT	TIME Rate per/hr	COST Rate per/hr	1.25 TOTAL TIME	1.25 TOTAL COST
EXTRA cards w/o tracings	Retrieve in public catalog	3,000 cards	30	$3	125	$ 375
*PREPARE copy of interim cataloged entries	REPRODUCE extra catalog card	12,000 cards				$ 600
**NUMBER & REPRODUCE non-MARC entries	FULL procedure	7,500 sheets	0.022736 entry	$0.115524 entry	171	$ 867
				INTERIM TOTAL:	296 hr	$ 1,842
				RETROSPECTIVE & INTERIM TOTAL:	5980 hr	$30,723

*Labor and material costs allocated from conversion (at 5¢ /card) for extra work done during regular production. No extra time allocated in this procedure.

**Estimated 50% non-MARC of 15,000 interim cards. Restatement of cost and time as:

Cost per unit (12¢, or $0.115524 per title)
Time per unit (82 secs, or .022736 hrs. per title)

FORMULAE USED IN PRODUCTION AND CONTROL

Phase A1: Initial Conversion

(Values of volumes and rates are defined below)

JOB	STEP	VOLUME/UNIT	TIME RATE per/hr	COST RATE per/hr	1.25* TOTAL TIME	1.25* TOTAL COST
1. NUMBER all catalog cards	Set up Number	500 runs C	5 runs 15000	N+K K		
2. SELECT main entries	Select Verify	C C-M	500 500	K L		
3. TRANSPORT cards		500 runs	5 runs	N+K		
4. NUMBER main entries, front and back	Set up Number	2M M	15000 200	K X+K		
5. REPRODUCE main entry fronts		250 runs	5 runs	K	See general equations in notes	
6. SELECT verso cards	Select Verify	M M-V	500 2000	K K		
7. COUNT verso cards	Set up Count	250 runs V	5 runs 15000	N+K K		
8. REPRODUCE verso cards		V	200	X+K		
9. INTERFILE verso sheets		V	700	K		
10. PACKAGE & label sheets		M+V	20 pkg= 10000 sheets	K		
11. REFILE main entry cards		M	700	K		
12. COUNT refiled cards cards	Set up Count	250 runs C	5 runs 15000	N+K K		

$$* \text{ 1.25 Total Time} = \frac{(1.25)(\text{Volume})}{\text{Time Rate}}$$

$$* \text{ 1.25 Total Cost} = \frac{(1.25)(\text{Volume})(\text{Cost Rate})}{\text{Time Rate}}$$

Assumed volumes and rates on next page.

VOLUMES AND RATES ASSUMED IN PRODUCTION AND CONTROL

C = 500,000 catalog cards.

M = 250,000 retrospective main entries; 15,000 interim.

V = 55,000 versos.

K = Clerical labor at $3/hr.

L = Librarian labor at $4/hr.

X = Xerox machine cost + cost per page of paper
= ($0.025 per copy) + ($0.0024 per page)
= $5.48/hr. (at rate of 200 copies per hour)

N = Numbering machine cost at $3000, distributed over 1500 runs
= $2/hr.

Notes: Serial main entries:

When selecting monograph main entries, Step 2, before numbering main entries, the estimated 6,000 serial main entries in the Official file, can be separately selected, copied, and stored until serial conversion begins. Under this approach, new serial main entries would be made from that time, and stored with the original 6,000. The "holdings" information would be converted, via reference to these original plus new main entries, to wherever "holdings" are recorded.

Alternative to selection at this point would be separate selection of 6,000+ from the 500,000+ file at time of serial conversion.

In either case, serial main entries must be separated from monograph main entries.

ILR Numbering experience:

Two full time clerks numbered and boxed 1,074,395 catalog cards, in 122 hours, at an effective rate of 11,000/hr. including 25% delay factor. Two clerks are advantageous, one to prepare, the other to number, in order to maintain continuous production. Accuracy of results was within .1% perfect. Preset control and counter recorded every card that was printed. Errors were due to skipped cards, or double feed, due to machine or type of input cards. Feed rollers need initial and set-up cleaning, and replacement every 150-200,000 cards. Mixed weights, paired, perforated, or otherwise linked cards required special handling or machine adjustment. Overprinting of numbers also occurs when cards have data where number is, are special brief listings, title page cards, withdrawn or marked over cards, marked duplicates. Skipped or overprinted numbers can be noticed during keying, noted where they occur in the deck, and recovered by reference to adjacent numbers.

SUMMARY: MONOGRAPHS--EDITING (Initial Conversion)

This Section A-2 presents estimated half-time costs for:

Two senior and eight regular editors to:

Edit reproduced sheets

Encode formatting and special data

Cross verify quality of work

One clerk to maintain control over filing, retrieval, and distribution of sheets.

Costs are based on:

(1) Three weeks of training.
(2) Four weeks of preliminary production.
(3) Seventy-one weeks of full production.

ESTIMATED TIME AND COST FOR EDITING 265,000 RECORDS

Phase A2: Initial Conversion

DESCRIPTION	PRODUCTION RATES	1.25 COST Rate	COST
Clerk (half time) file, control, and retrieve sheets	18 mos.	$ 6,000/yr.	$ 4,500
* Senior Editor (half time)(2) edit, coordinate, and control procedures	21 mos.	$ 9,000/yr.	$15,750
Editors (half time)(8) edit and encode sheets:	18 mos. = 1500 hrs. (250 work day yr.)	$ 8,000/yr.	$48,000
I. 1st week — train in MARC format, coding manual, and library cards.	(20 hrs. training)		
2&3rd weeks — Edit sheets (Supervisor cross checks each actual record)	(1.25)(4.5 min.) =5.625 min./record =142 avg. records/wk. of 1st 3 weeks		
II. 4-7th weeks — Edit sheets Cross check 10% records for group uniformity. (Supervisor cross checks 2% of records)	(1.25)(3.5 min.) (.1)(1.25)(3 min.) =4.75 min./record =248 records/wk. of next 4 weeks		
III. 71 weeks — Edit sheets Cross check 2%	(1.25)(2.5 min.) (.02)(1.25)(3 min.) =3.2 min./record =370 records/wk.		
	27,264 records/18 mos per Editor.		
	Max. 272,264 records	Total	$ 68,250

* Senior Editors chosen before regular editors for preparation in procedures.
Rates based on ILR editing of 11,227 sheets.
Unit Cost = $.26/entry ($ 68,250/265,000 entries)
Unit Rate = avg. 350 records/week

SUMMARY: MONOGRAPHS--KEYING (Initial Conversion)

This Section A-3 presents estimated costs for:

Keypunching edited sheets onto tab cards, in a continuous string format (record number--card number--data).

Costs are based on two rates:

(1) Preliminary (allowing training, slower production, and higher unit cost, although actual data is involved).

(2) Full production.

Although estimates are stated for keypunching, the continuous string format is compatible with scanner typewriters or on-line consoles, whenever these become economically feasible.

2

ESTIMATED TIME AND COST USED IN KEYING

Phase A3: Initial Conversion

DESCRIPTION	PRODUCTION	VOLUME (Tab Cards)	1.25 COST Rate	COST
First 9 wks. of editing will backlog 21,610 records = (108,050 cards). Thereafter, 3700 records = (18,500 cards) are available per week.	Preliminary	10,000	$86.50/ 1,000	$ 865
	Full	1,315,000	$70.00/ 1,000	$92,050
Material (cards) @1.40/1000 cards				$ 1,855
			TOTAL:	$94,770

Assume outside firm, or State keying pool. Keypunch assumed, though format is compatible with scanner typewriters, or on-line consoles.

Keying of MARC partial identification assumes 1 card. If non-MARC, rest of data added in cards. Initial 12000 interim will be test of MARC procedure, and probably full keying in many of them. Estimate includes all 12000 interim + 3000 (without tracing) as test of on-CSL-tape procedure.

Average number of characters per record = 350. 80 tab chars/ tab card = 5 tab cards/record.

Unit cost = $.0715/tab card, or $.3575/entry.

SUMMARY: MONOGRAPHS--COMPUTER PROCESSING (Initial Conversion)

This section A-4 presents estimated costs for:

Computer processing of:

Original records.

Corrections.

Merging approved records into established Index, Hold, and Bibliographical Master files.

Costs are based on:

(1) Computer reading of input tab cards onto magnetic tapes.

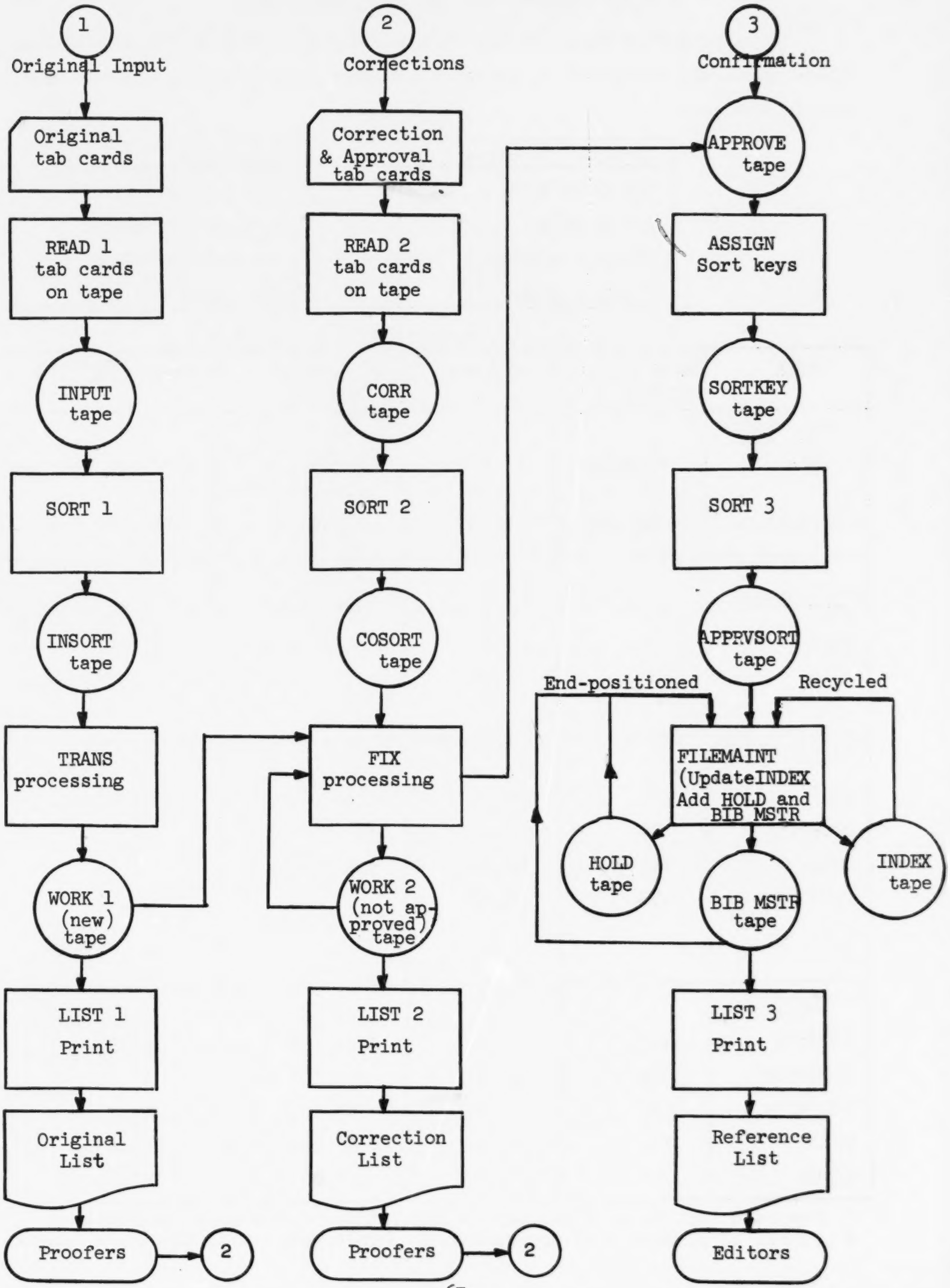
(2) Processing of original records.

(3) Processing of correction records.

(4) Computer merging of approved records.

(5) Computer printing of output for proofing.

INITIAL COMPUTER PROCESSING
DETAIL INPUT/OUTPUT AND PROGRAMS
(Adapted from SYS Fig. 11: PROCESSING CYCLES FOR CONVERTING CATALOG RECORD)
Vol. 1, p.49



ESTIMATED TIME AND COSTS USED IN COMPUTER PROCESSING

Phase A-4: Initial Conversion

Costs projected here concern five basic phases of computer processing. These phases are estimated in terms of assuming the following machine speeds and file sizes:

RATES

Computer Processing	Rate per unit
Tab Card Read	27,000 tab cards per hour
I/O limited	10 minutes per tape
Compute limited	25 minutes per tape
Printer speed	20,000 lines per hour

FILE SIZES

FILE	CHARS. PER RECORD	RECORDS PER TAPE	RECORDS PER FILE	TAPES PER FILE
(1)TAB CARDS				
Original	400/rec.	265,000 entries	5 cards/ entry	1,325,000 tab cards
Corrections	80/line	*204,000 lines	1 card/ line	204,000 tab cards
TAPES				
(2)Original				
INPUT	400	25,000	265,000	10.6 tapes
INSORT	400	25,000	265,000	10.6
WORK 1	500	20,000	265,000	13.25
(3)Corrections				
CORR	100	125,000	*204,000	1.63
COSORT	100	125,000	*204,000	1.63
WORK 2	500	20,000	*163,100	8.16
APPROVE	500	20,000	265,000	13.25
(4)Approved				
SORTKEY	500	20,000	265,000	13.25
APPRVSORT	500	20,000	265,000	13.25
BIB MSTR	500	20,000	265,000	13.25
HOLD	250	40,000	#265,000	6.63
INDEX	50	200,000	265,000	1.33

* See next section on proofing for derivation of number of corrections.
\# Number of records will approximately double when other libraries added.

FILE	CHARS. PER LINE	LINES PER RECORD	RECORDS PER FILE	LINES PER FILE
(5) PRINT LINES				
Original (Full)	132	35 lines/rec. 2 recs./line	265,000	4,637,500
Corrections	132	5 lines/rec. 2 recs./line	204,000	510,000
Controls	75 runs	200 lines/run	--	15,000
Final Format Master	132	10 lines/rec. 2 recs./line	265,000	1,325,000

Printing is assumed to include skips, and a print chain with upper and lower cases, and special diacritic marks.

Approvals are assumed to be entered along with correction cards as implied non-entries of batch, or individual or inclusive series of numbers.

Computer processing is assumed to start the 12th week after training and preliminary editing and keying have accumulated a backlog. This means approximately 4,000 records processed weekly.

Computer cost is assumed at $100 per hour.

In terms of these rates and file sizes, the chart which follows shows a breakdown of estimated costs for the computer processing during initial conversion.

Computer processing is described as a type of input processed by a type of program to produce a type of output. For example, Step 2a "INSORT processed by TRANS to WORK 1" describes "Input (INSORT) processed by program (TRANS) to produce output (WORK 1)"--parts extracted for descripion underlined.

In all programs the time for reading input and writing output tapes overlap. Hence the total time (or number of tapes) represents the larger number of the two, rather than their sum.

Programs which could be processed continuously in time and order are described as separate steps for completion, rather than as continuous processing internally linked within the computer. Such separations avoid extra computer programming needed to achieve such linkage, secure back-up for the various steps of processing, and allow the computer center to have options in actual scheduling.

INITIAL COMPUTER PROCESSING

PROCESSING	RATE	NO. OF UNITS	TIME	COST
(1) Tab Cards				
a. Orig. tab cards put on tape by READ 1 to INPUT	Card	1,325,000 cards	49.07 hrs	$ 4,907
b. Corr. tab cards put on tape by READ 2 to CORR	Card	204,000	7.56	756
(2) Original				
a. INPUT sorted by SORT 1 to INSORT	*Sort 1	10.60 tapes	14.16	1,416
b. INSORT processed by TRANS to WORK 1	Compute	13.25	5.52	552
(3) Corrections				
a. CORR sorted by SORT 2 to COSORT	*Sort 2	1.63	2.12	212
b. COSORT, WORK 1 and WORK 2 processed by FIX to APPROVE and WORK 2	Compute	23.04	9.60	960
(4) Approved				
a. APPROVE processed by ASSIGN to SORTKEY	I/O	13.25	2.21	221
b. SORTKEY sorted by SORT 3 to APPRVSORT	*Sort 3	13.25	17.70	1,770
c. APPRVSORT and INDEX, & end-positioned BIB MSTR and HOLD, processed by FILEMAINT to INDEX, and added HOLD and BIB MSTR	I/O	21.21	3.54	354
(5) Printed Lists				
a. WORK 1 printed by LIST 1 to Original-List	Print	4,652,500 lines	232.63	23,263
b. WORK 2 printed by LIST 2 to Correction-List	Print	525,000	26.25	2,625
c. BIB MSTR printed by LIST 3 to Reference-List	Print	1,325,000	66.25	6,625
	Sub Total Operations:		436.61 hrs	$43,661
Tape reels purchased(18mos) @ $20/reel, 10/month	Sub Total Materials:			3,600
	Total Time/Cost		436.61 hrs	$47,261

Unit Time= 6 secs/record=(436.61)(3600)/265000

Unit Cost= $0.18/record =$47261/265000

*** Basis of Sort rates given on following page.**

BASIS OF SORT RATES

Let:

S = number of strings within the sort = (RC/A)

b = order of merge in sort

$[\log_b S]$ = high integer log base b of S = number of passes per sort

R = records per run

C = characters per record

A = allotted characters for sort in core

F = records per file

H = hourly rate per tape

T = records per tape

Sort hours* = $[\log_b (RC/A)](HF/T)$

or stated verbally:

$$S = \frac{(\text{records per run})(\text{characters per record})}{(\text{allotted characters for sort in core})}$$

$$\text{Sort hours}^* = [\log_b S]\frac{(\text{hourly rate per tape})(\text{records per file})}{(\text{records per tape})}$$

Assume:

b = 2 order of merge in sort

R = 4000 records per run

A = 10000 allotted characters for sort in core

H = 0.167 hourly rate per tape = 10mins/tape

Then:

Sort 1 = $[\log_2(4000(400)/10000)](0.167)(265000)/25000$

= 14.16 hrs

Sort 2 = $[\log_2(4000(100)/10000)](0.167)(265000)/125000$

= **2.12** hrs

Sort 3 = $[\log_2(4000(500)/10000)](0.167)(265000)/20000$

= 17.70 hrs

* Derived as: Sort hours = (Time per sort)(Number of sorts)

Time per sort = (Number of passes per sort)(Proportion of records per run per tape)(Hourly rate per tape)

= $[\log_b S]$ x (R/T) x (H)

Number of sorts = (records per file) / (records per run)

= F / R

Combining and simplifying:

Sort hours = $[\log_b S]\frac{R}{T}(H)\frac{F}{R} = [\log_b S](HF/T)$ or $[\log_b (RC/A)](HF/T)$

SUMMARY: MONOGRAPHS--PROOFING (Initial Conversion)

This section A-5 presents estimated half-time costs for:

Two senior and eight regular proofers to:

Scan computer output for:

(1) MARC formatting.

(2) Content.

Prepare corrections by:

(3) Deciding correction strategy.

(4) Writing corrections.

Rescan corrections on computer output for:

(5) MARC formatting.

(6) Content.

Cross verify quality of work.

One clerk to maintain control over filing, retrieval, and distribution of sheets.

Costs are based on:

(1) Three weeks of <u>training</u>.

(2) Four weeks of <u>preliminary</u> production.

(3) Seventy-one weeks of <u>full</u> production.

ESTIMATED TIME AND COSTS USED IN PROOFING

Phase A:5 Initial Conversion (265,000 records)

Costs projected here concern three basic phases of proofing: training, and preliminary and full production. These phases are estimated in terms of assuming the following steps of operation, file sizes, and rates of proofing.

S T E P S	*FILE SIZE Sum of cycles		C Y C L E S			
		Orig. cycle	Recycles 1	2	3	4
1. Scan record format	265,000	265,000	-	-	-	-
2. Scan record content	265,000	265,000	-	-	-	-
3. Decide correction strategy	163,078	159,000	3,975	100	3	-
4. Write correction lines	203,847	198,750	4,968	125	#4	-
5. Scan corr. line format	203,847	-	198,750	4,968	125	#4
6. Scan corr. line content	203,847	-	198,750	4,968	125	#4

* Basis of File Sizes on the following page.
\# Last 4 corrections are assumed not to require any re-entry.
In general, one cycle is assumed as "entry-of-correction + re-scan".
Hence, a maximum of four cycles is involved here.

Approvals are assumed to be entered implicitly by omission of corrections within a batch. If schedule cannot be maintained for entering all corrections within a batch just processed, separate approval cards can be re-entered by individual or inclusive series of numbers.

Proofing is assumed for all records. Only a given percentage of records, however, may be proofed at different stages, according to priority of purposes:

(a) Monitoring the quality of keying.

(b) Verifying critical fields.

(c) Sampling a statistically reliable percentage as a guide to how much further proofing is required to achieve an acceptable level of accuracy.

BASIS OF FILE SIZES

Let:

R = number of records

M = number of MARC lines (or data elements) per record

P = percentage of error

N = number of lines per correction

k = number of correction cycles (1,2,...,k)
cycle defined as "entry-of-correction + rescan)

CORRECTIONS AND CORRECTION-LINES

C_k = number of corrections in cycle k

$$= \left[\begin{matrix}\text{number of lines}\\ \text{per correction}\end{matrix}\right]^{k-1} \times \left[\begin{matrix}\text{percentage}\\ \text{of error}\end{matrix}\right]^{k} \times \left[\begin{matrix}\text{number of}\\ \text{MARC lines}\end{matrix}\right] \times \left[\begin{matrix}\text{number of}\\ \text{records}\end{matrix}\right]$$

$$= (N^{k-1}) \times (P^{k}) \times (M) \times (R)$$

L_k = number of correction-lines in cycle k

$$= \left[\begin{matrix}\text{Number of}\\ \text{correction-lines}\end{matrix}\right] \times \left[\begin{matrix}\text{Number of corrections}\\ \text{in cycle k}\end{matrix}\right]$$

$$= (N) \times (N^{k-1}P^{k}MR)$$

SUM OF CORRECTIONS AND CORRECTION-LINES

ΣC_k = sum of cycles of corrections

$$= C_1 + C_2 + \ldots + C_k$$

$$= N^0P^1MR + N^1P^2MR + \ldots + N^{k-1}P^kMR$$

$$= PMR(1 + NP + \ldots + N^{k-1}P^{k-1})$$

$$= PMR(1 + NP + \ldots + N^{k-1}P^{k-1})(NP - 1)/(NP - 1)$$

$$= PMR\,(N^kP^k - 1) / (NP - 1)$$

ΣL_k = sum of cycles of correction-lines

$$= \left[\begin{matrix}\text{Number of}\\ \text{correction-lines}\end{matrix}\right] \times \left[\begin{matrix}\text{Sum of cycles}\\ \text{of corrections}\end{matrix}\right]$$

$$= (N) \times (\Sigma C_k)$$

$$= N(PMR)(N^kP^k - 1)/(NP - 1)$$

Assumed-Values

R = 265,000 records

M = 30 MARC lines per record (to emphasize format)

P = 0.02 (2% error)

N = 1.25 lines per correction (75% 1 line; 25% 2 lines)

k = 4 cycles of corrections

Corrections

Then:

$$\Sigma C_k = PMR(N^k P^k - 1) / (NP - 1)$$

$$= (0.02)(30)(265000)[(1.25)^4(0.02)^4 - 1] / [(1.25)(0.02) - 1]$$

$$= 163,078$$

Correction-lines

$$\Sigma L_k = N(\Sigma C_k)$$

$$= (1.25)(163,078)$$

$$= 203,847$$

RATES OF PROOFING

	PHASES OF PRODUCTION		
	Training	Preliminary	Full
Unit Rate per Record in each phase	343.270 sec. = 5.721 min.	244.712 sec. = 4.079 min.	157.693 sec. = 2.628 min.

BASIS OF PROOFING RATES

Effective Rates for steps "i" in phases "j" :

Let:

f_i = file size at each step i

$r_{i,j}$ = regular proofing rate at step i in phase j
where $1 \leq i \leq 6$ and $1 \leq j \leq 3$

d = delay factor

R = number of records

$E_{i,j}$ = effective rate of proofing at step i in phase j

= [regular + delay] [regular rate] [proportion of file size to number of records]

$= (1 + d)(r_{i,j})(f_i / R)$

Unit Rates in terms of effective rates;

U_j = sum of effective rates for steps in each phase

$$= \sum_{i=1}^{6} (1 + d)(r_{i,j})(f_i / R)$$

$$= \left[\frac{(1 + d)}{R}(r_{i,j}f_1 + r_{2,j}f_2 + \ldots + r_{6,j}f_6)\right]$$

BASIS OF PROOFING RATES (cont.)

Assumed Values

Step	FILE SIZES = f			
f_1 =	265,000			
f_2 =	265,000			
f_3 =	163,078			
f_4 =	203,847			
f_5 =	203,847			
f_6 =	203,847			
Step	REGULAR RATES			
	Phase 1	Phase 2	Phase 3	
$r_{1,j}$ =	90 sec.	60 sec,	30 sec.	
$r_{2,j}$ =	90	75	60	
$r_{3,j}$ =	60	30	15	
$r_{4,j}$ =	15	15	15	
$r_{5,j}$ =	30	20	10	
$r_{6,j}$ =	30	20	10	
d =	0.25 (25% delay factor)			
R =	265,000 records			

Then: (Simplifying before calculation for common file sizes):

$$U_j = \left[(1 + d)(r_{1,j} + r_{2,j} + \frac{r_{3,j} f_3 + (r_{4,j} + r_{5,j} + r_{6,j})(f_4)}{R}\right]$$

$$U_{j=1} = (1 + 0.25)(90 + 90 + \frac{60(163078) + (15 + 30 + 30)(203847)}{265000}$$

$$= 343.270 \text{ sec.}$$

$$U_{j=2} = (1 + 0.25)(60 + 75 + \frac{30(163078 + (15 + 20 + 20)(203847)}{265000}$$

$$= 244.712 \text{ sec.}$$

$$U_{j=3} = (1 + 0.25)(30 + 60 + \frac{15(163078) + (15 + 10 + 10)(203847)}{265000}$$

$$= 157.693 \text{ sec.}$$

ESTIMATED TIME AND COST FOR PROOFING 265,000 RECORDS

Phase A5: Initial Conversion

DESCRIPTION	PRODUCTION RATES	1.25 COST Rate	COST
Clerk (half time) file, control, and retrieve sheets	18 mos.	$6,000/yr.	$ 4,500
*Senior Proofers (half time) (2) proof, coordinate and control procedures	21 mos.	$9,000/yr.	$ 15,750
Proofers (half time) (8) scan sheets and make corrections:	18 mos. = 1500 hrs. (250 work day yr.)	$8,000/yr.	$ 48,000
1. 1st week — Train in MARC format, proofing manual, and print-out sheets	(20 hrs. training)		
2&3rd weeks — Scan sheets & make corrections (Supervisor cross checks each actual record)	5.721 min./record =140 avg. records/wk. of 1st 3 weeks		
II. 4-7th weeks — Scan sheets & make corrections. Cross check 10% of records for group uniformity. (Supervisor cross checks 2% of records)	(1.1)(4.079 min/recs = 4.5 min./rec = 267 recs/wk. of next 4 weeks		
III. 71 weeks — Scan sheets & make corrections. Cross check 2% of records	(1.02)(2.628 min/rec = 2.7 min/rec = 445 recs/wk.		
	33,083 records/18 mos per proofer.		
	#Max. 330,830 records	Total	$ 68,250

*Senior Proofers chosen for preparation in procedures before regular proofers. Rates based on ILR proofing.
Unit Cost = $.26/entry ($ 68,250/265,000 entries)
Unit Rate = 424 records/week = 2.83 mins./record. (See next comment)

Extra time not needed for proofing monographs can be used in proofing serials. Any estimate, however, for proofing must allow for unexpected increases in number of corrections. Maximum feasible 330,830 records vs. 265,000 expected records provides for this.

SUMMARY: SERIALS--Complete Initial Conversion

This section A-6 presents proportionately estimated costs covering: Complete initial conversion of serials in successive phases of:

RETROSPECTIVE

(1) Production and Control (number and xerox source records)

(2) Editing (select, transcribe, and encode information on one sheet)

(3) Keying (key edited sheets in continuous string format)

(4) Computer Processing (enter and correct records to establish files)

(5) Proofing (verify formats and data, and recycle corrections)

INTERIM

(6) Enter new records via MARC serials(if available) or as full conversion.

(7) Enter interim holdings via turn-around cards, verified directly from updated source records.

Costs are based on estimates derived from equivalent amounts of data and similar processing established as costs for converting monographs.

ESTIMATED UNIT COSTS FOR INITIAL CONVERSION OF SERIALS

Steps for converting serials were summarized parallel to monographs (Introductory pages 13/14, this volume). An overall approach to conversion is described in Vol. 4, pages 171-180, covering three types of effort, called Conversion Levels (CONVEL I, II, and III):

I. Convert the basic identifying data (similar to monographs), aided where possible by Automatic Field Recognition (AFR) programs.

II. Complete and correct this identification data, and add data unique to serials: full history of possible issues, current publication pattern, and cross references to other serials in its history, not including more history than is held by at least one library, except perhaps for continuity. This full description of any serial new to the file may be more complete than the actual holdings in any particular library. The full record is maintained on the Central Master File (CMF).

III. Convert actual holdings and local variations for particular library, maintained on the Local Master File (LMF).

Estimate of unit cost for these three levels of conversion effort is derived per record from the amount of data, and the amount of processing (both manual and computer) in terms of monograph units of conversion effort.

Type of Record	Avg.Amt. of Data	Avg. Amt. of Processing
Monographs	400 chars.(1:1)	1:1
Serials	1600 chars.(4:1)	* 1.1:1

* reflecting the comparatively more complex:
- manual effort in collecting, selecting, and transcribing conversion information from more than one source document, at each level;
- computer processing involving more intricate logic and character-for-character searching within a computer record for each issue-item.

Unit costs for Initial Conversion of Serials (new to file) calculated as:

No. of Data Monog.Units	x	No. of Processing Monog. Units	x	Unit Cost of D/P Monog.Units	=	Avg. Unit Cost-Serial
4	x	1.1	x	$1.18	=	$ 5.19

The manual aspect of this much higher estimated unit cost for serials is specifically provided for on page 28 by additional personnel, plus overlap in work by production clerks on both monographs and serials.

SUMMARY: CONVERSION OF OTHER LIBRARIES

This section B presents estimated costs for:

(1) Conversion of other Monographs.

Costs for these conversions are estimated on basis of:

Unit prices required for full conversion of entries.

Unit prices required for partial converisons, using established CSL/PC or MARC files.

(2) Conversion of other serials.

B. CONVERSION OF OTHER LIBRARIES - -(1) MONOGRAPHS

This section covers the conversion of the retrospective holding of other libraries; "Other" here simply means other than the first library to be converted. The distinguishing condition is the existence of a Bibliographic Master file which is assumed to contain 50% of the holdings of the library to be converted. Each of the costs incurred under initial conversion will be discussed separately, both as to unit costs and total volumes projected over three years. Since the first conversion is scheduled to take 18 months, in essence the conversion of other libraries is limited to an 18 month projection during the second and third year of the Center's operation.

The volume of conversion work during the first 18 months is 265,000 entries. However during conversion of other libraries, the basic labor requirement changes in that 50% of the new entries are already in the Master file and require only partial editing (one-third as much) and no proofing. Thus, assuming a 50% duplication rate among new entries, the edit-proofing rate rises from 70 per day (during initial conversion) to 120 conversions per editor-day (including both full edit-proof and partial search-edit conversions). Thus the weekly rate is approximately 600 per week, or 6,000 per week for the 10 man editorial staff. The 18 month projected volume, given that the 10 man staff remains constant, is therefore 450,000.

1. Production and Control

   a. Unit Cost - .12 (Same as Initial Conversion)

   b.

Volume	2nd Yr.	3rd Yr.	Total
cards:	150,000	300,000	450,000
cost:	$ 18,000	$ 36,000	$ 54,000

2. Editing

   a. Unit Cost: for records not in the Master File, the unit cost remains $.26; the search-edit cost for records not in the file is estimated at $.09. A combined unit cost is therefore $.18.

b.

Volume:	2nd Yr	3rd Yr	Total
cards (full)	75,000	150,000	225,000
cost	$19,500	$39,000	$58,500
cards (partial)	75,000	150,000	225,000
cost	$ 6,750	$ 13,500	$20,250
cards (all)	150,000	300,000	450,000
cost	$26,250	$52,500	$ 78,750

3. Keying

a. Unit Cost: full records @$.35; partial records @$.12.

b.

Volume:	2nd Yr	3rd Yr	Total
cards (full)	75,000	150,000	225,000
cost	$26,250	$52,500	$78,750
cards (partial)	75,000	150,000	225,000
cost	$9,000	$18,000	$27,000
cards (all)	150,000	300,000	450,000
cost	$35,250	$70,500	$105,750

4. Computer Processing

a. Unit Cost: for full records, new to the file, the unit cost has been estimated at $.18 per record, or 6 secs. The time and cost for records already in file is significantly less, because of reduced record size and no need to print Bibliographic Master Records. Thus the processing time is given as 1 second per record, or $.03.

b.

Volume	2nd Yr	3rd Yr	Total
records (full)	75,000	150,000	225,000
time	125 hrs	250 hrs	375 hrs
cost	$12,500	$25,000	$37,500
records (partial)	75,000	150,000	225,000
time	21 hrs	42 hrs	63 hrs
cost	$2,100	$4,200	$6,300
records (all)	150,000	300,000	450,000
time	146 hrs	292 hrs	438 hrs
cost	$14,600	$29,200	$43,800

5. Proofing

a. Unit Cost: For full records the unit cost is $.26; the cost incurred for partial records is $.04.

b. Volume	2nd Yr.	3rd Yr.	Total
records (full)	75,000	150,000	225,000
cost	$ 19,500	$ 39,000	$ 58,500
records (partial)	75,000	150,000	225,000
cost	$ 3,000	$ 6,000	$ 9,000
records (all)	150,000	300,000	450,000
cost	$ 22,500	$ 45,000	$ 67,500

6. Verification

a. Unit Cost: $.11/.07 per unique entry to file. Unique here is defined as first: a title not already duplicated in the Master file, and second: a record whose names or subject terms are not already duplicated.

b. To help understand the authority verification costs, it would be useful to re-examine the total authority verification volume projection, both for initial and other conversions.

Category	YEAR 1	2	3	Total
"initial entries"	180,000	85,000	0	265,000
number unique	180,000	85,000	0	265,000
cost	19,800	$ 5,950	0	$ 25,750
other entries (full)	0	75,000	150,000	225,000
number unique	0	37,500	75,000	112,550
cost	0	$ 2,625	$ 5,250	$ 7,875
"other" entries (partial)	0	75,000	150,000	225,000
number unique	0	0	0	0
cost	0	0	0	0
total entries verified	180,000	122,500	75,000	377,500
total cost	19,800	$ 8,575	$ 5,250	$ 33,625

Note that these figures can be related to the growth of the system files BIB MSTR, BIB MSTR, HOLD, and INDEX.

	INDEX	BIB MSTR	HOLD
Year 1	175,000	175,000	175,000
Year 2	340,000	340,000	415,000
Year 3	490,000	490,000	715,000

B. CONVERSION OF OTHER LIBRARIES--(2) SERIALS

Converting serials in other libraries will cost more than converting monographs, even though title information is already on tape. This greater cost is due to additional information essential to serials, such as local-holdings, operational information, and possible cross references to former titles or other serials.

Unlike monographs, serials also involve continuous adding of new information, again issue data, cross references, or items pertaining to operations for a particular library. This further updating, which closely resembles conversion, will continue as long as any issues are ever acquired to replace losses, fill in gaps, or maintain the full run and any changes announced by the publisher of a serial.

As far as computer processing of serials is concerned, then, to replace or acquire one or many issues of the same title is no different from converting one or many issues of the same title initially.

Accordingly, further steps for such interrelated conversion and maintenance of serials are estimated here together. As in Initial Conversion, estimated unit costs will be derived from the amount of data and the amount of processing (both manual and computer) in terms of monographic units of conversion effort, with special breakdown in the maintenance.

Assumed volumes: (V) 1st yr. 500 new; 2nd yr. 4900(3800 new, 600 dup, 500 previous)
3rd yr. 18000 (3700 new, 9400 dup, 4900 previous)

STEP	BASIS OF COST	UNIT COST	COSTS PER YEAR 1	2	3
B. Conversion of other libraries Dup	(3)(1.1)($1.18)	$1.32	-	$ 790	$12,410
New	(as in Initial)	$5.19	-	$1,555	$ 8,825
C. Maintenance:	(in terms of V)				
Reg.:Weekly Update CPU					
Monthly Extracts pro-	.11583 V		$ 58	$ 568	$ 2,085
Other:Local & Union cess-		1st yr. $1.088			
Active, On-Order, ing.					
Profile,Inaccessible					
Sort Processing: Regular and Other (1.34 Regular)	.000957912Vhrs. at $100/hr. (cf.p.71)	2nd yr. $0.485	$113	$1,100	$ 4,037
Print Arrivals(.65V),Claims (.15V),Binding(.08V(,Pyt.(1.5V)		3rd yr. $0.433			
New Orders (.01V) Monthly	.0239V				
Other	.032V		$ 28	$ 274	$ 1,006
Punch & Interpret Regulars	2.15V/150 + 2.15V/3000		$ 9	$ 75	$ 270
Set up each run at $5/run			$336	$ 360	$ 390
TOTAL MAINTENANCE:			$544	$2,377	$ 7,788
D. Authority Verification: New	(as in Mono-	$0.48	$240	$1,970	$ 1,775
Dup	graphs)	$0.31	-	$ 95	$ 2,915
E. Data Retrieval	(as in Mono-graphs)	$0.09	$ 45	$ 395	$ 1,180

SUMMARY: MAINTENANCE OF ESTABLISHED FILES

This Section C presents estimated costs for daily runs to process records received from the Library of Congress MARC distribution service. The processing consists of merging MARC records into the file, processing search requests against the MARC file, and printing out catalog cards.

(Since costs for this last service will be allocated to the search requestor, it is not included here.)

C. MAINTENANCE OF ESTABLISHED FILES (MARC)

The cost here has to do with (a) incorporating MARC records into the system files; and (b) searching MARC records against requests in support of current cataloging. No cost estimate is made here of the volume of requests, nor of the total printing cost.*

The operating strategy here is to perform a daily run against the MARC file. The annual MARC input is estimated at 60, 70, and 80 thousand records over the next three years. Assuming half of those file sizes as an estimating parameter, and 20,000 records as a tape packing factor, the following table can be constructed.

	YEAR		
	1	2	3
File Size ÷ 2	30,000	65,000	105,000
No. of tapes	1.5	3	5
Daily processing time	.25 hr.	.5 hr.	.83 hr.
Daily cost	$25.	$50.	$83.
Annual time	62.5 hrs.	125 hrs.	207.5 hrs.
Annual cost	$6250	$12,500	$20,750
Cumulative file size	60,000	130,000	210,000
Unit cost	$.10	$.10	$.10

---

*The following unit costs can be computed: Annual MARC File ÷ 250 days = 280 records per day; at 10 lines per record = 2800 print lines; X 4 cards = 11,200 lines; @ 20,000 lines per hour = .56 hrs. = $56.00. Unit cost may thus be projected as $.20 per record for 4 copies of each catalog card, or $.05 per card.

III - D

SUMMARY: AUTHORITY VERIFICATION

This section D presents estimated costs for:

Establishing and maintaining authority files of authors and subjects.

Updating corresponding authors and subjects on master files.

These costs would involve:

Manual processing.

Machine processing of:

Master files (BIB MSTR, HOLD, and INDEX)
Transaction files (SUBJECT and NAMES)
Output files (CONSOLIDATED)

D. AUTHORITY VERIFICATION

The cost being projected here concerns the verification of author names and subject headings used in the three Processing Center files: BIB MSTR, INDEX, HOLD. There are fifteen basic processing (manual and machine) steps required in the process. Before elucidating the cost per processing step, it is necessary to set up some assumed file sizes, growth rates and machine speeds.

As in section A we will discuss processing time in terms of tapes:

I/O limited: .15 hr per tape

compute limited: .4 hr per tape

Printer speed: 20,000 lines per hour

FILE SIZES OF RECORDS TO BE VERIFIED

		Year 1		Year 2		Year 3	
FILE	RECS/ TAPE	RECS	TAPES	RECS	TAPES	RECS	TAPES
BIB MSTR	20k	180,000	(9.)	127,500	(7.)	75,000	(4.)
SUBJ TRANS	100k	270,000	(2.7)	191,250	(2.)	112,500	(1.1)
NAME TRANS	100k	270,000	(2.7)	191,250	(2.)	112,500	(1.1)
SUBJ CONSOL	20k	40,000	(2.)	28,500	(1.5)	16,825	(1.)
NAME CONSOL	55k	135,000	(2.5)	95,625	(2.)	56,250	(1.)
SUBJ PROOF	20k	10,000	(.5)	7,125	(3.5)	4,200	(.2)
NAME PROOF	40k	135,000	(3.3)	47,800	(2.5)	28,125	(1.5)
SUBJ UPDATE	100k	35,000	(.3)	25,200	(.2)	14,700	(.2)
NAME UPDATE	100k	135,000	(1.3)	95,625	(1.)	56,250	(.5)

The file sizes above are based on estimating 1.5 subject headings per file record, .5 author tracings and 1 main entry author name. The above sizes and rates (machine time at $100.00) give all the assumptions required, with the exception of manual authority editing. This is taken to be at 1/3 the basic cost of editing a single catalog record, or $.08 (1.2 minutes) for both proofreading and correction.

	Processing		Yr. 1	Yr. 2	Yr. 3
1.	Pass Thru BIB MSTR		1.35 hrs.	1.05 hrs.	.6 hrs.
2.	Sort SUBJ TRANS		2.58	1.65	.88
3.	Sort NAMES TRANS		2.58	1.65	.88
4.	Consolidate SUBJ		.41	.3	.16
5.	Consolidate NAMES		.41	.3	.16
6.	Match against previous authority files		.68	.52	.30
7.	Print out PROOF		7.25	5.1	3.0
8.	Verify	(volume)	(145,000)	(55,000)	(32,400)
		(time)	(2733)	(1100)	(648)
		(cost) @$4/hr.	($ 11,600)	($4,400)	($2,592)
9.	Correct	(volume)	(75,000)	(27,500)	(16,000)
		(time)	(1500)	(550)	(320)
		(cost) @$4/hr.	($ 6,000)	($2,200)	($1,280)
10.	Generate UPDATE files		.64 hrs.	.48 hrs.	.28 hrs.
11.	Sort UPDATE		1.2	1.0	.45
12.	Update BIB MSTR		2.85	1.05	.6
13.	Sort INDEX into BRIDNO order		1.5	3.1	4.8
14.	Update INDEX		.26	.51	.74
15.	Resort INDEX		1.5	3.1	4.8
	Machine Time		23.21	19.81	17.65
	+ 20% rerun		2.32	1.98	1.76
			25.53 hrs.	21.89 hrs.	19.41 hrs.
	cost (machine)		$ 2,550	$ 2,190	$ 1,940
	cost (manual)		$ 17,600	$ 6,600	$ 3,872
		TOTAL	$ 20,150	$ 8,790	$ 5,812
	unit cost		$ .11	$ .07	$ .07

SUMMARY: DATA RETRIEVAL

This section E presents estimated costs for:

Consolidating holdings into the master record.
Constructing subfiles required to do this.

Costs are based on computer processing of established or new files:

(1) Select local holdings (HOLD file)
(2) Sort selections (TRANS file)
(3) Merge/select with master (BIB MSTR file)
(4) Sort consolidated records (BOOK file)

E.DATA RETRIEVAL

Data retrieval is defined in the current version of the system to mean construction of files of holdings for libraries, optionally partitioned by time periods. The processing logic basically consists of:

1. Selection Pass against HOLD file = TRANS file
2. Sort TRANS file into BRIDNO order
3. Merge-select pass against BIB MSTR = BOOK file
4. Sort BOOK into library order

The following file size estimates have been used:

		YEAR		
FILE	RECS/TAPE	1	2	3
INDEX	200,000	235,000	470,000	600,000
no. of tapes		2	3	4
BIB MSTR	20,000	235,000	470,000	600,000
no. of tapes		12	24	30
HOLD	40,000	175,000	415,000	715,000
no. of tapes		5	11	18

(The above estimates include MARC records.)

Processing costs are estimated here on the basis of monthly runs. It is assumed: (a) that 50% of the HOLD file will be retrieved each year, in 12 equal monthly portions; and (b) that the processing time for each year be based on the largest file size for that year, even though the file is growing regularly during the year toward its final size.

PROCESSING	Monthly	YEAR 1	2	3
1. pass against HOLD file	time:	.83 hrs.	1.83 hrs.	3.0 hrs.
	cost:	$83	$183	$300
2. Sort TRANS into BRIDNO order	size:	7000	17,000	40,000
	time:	.16	.33	.91
	cost:	$16	$33	$91
3. Merge against BIB MSTR	time:	2.0	4.0	5.0
	cost:	$200	$400	$500
4. sort Book	time:	.16	.33	.91
Monthly Total	time:	3.15	6.50	9.82
	cost:	$315	$650	$982
Annual Total	time:	37.8	78.0	117.8
	cost:	$3780	$7800	$11,780

APPENDIX IV: OUTPUT SERVICES

SUMMARY:

This appendix presents a general flow chart of five steps, followed by detail bases for costs to:

1. Generate added entries and sort keys, and to sort file

2. & 3. Convert tape to page image and master.

4. Reproduce and collate copies from master.

5. Bind and prepare finished catalogs.

OUTPUT SERVICES - CATALOGS
DETAILS OF ALTERNATIVE METHODS
(Adapted from SYS Fig. 14 A OUTPUT PROCESSING FLOW, Vol. 1, p. 56)

STEP

1. Generate added entries and sort-keys

   Sort full file.

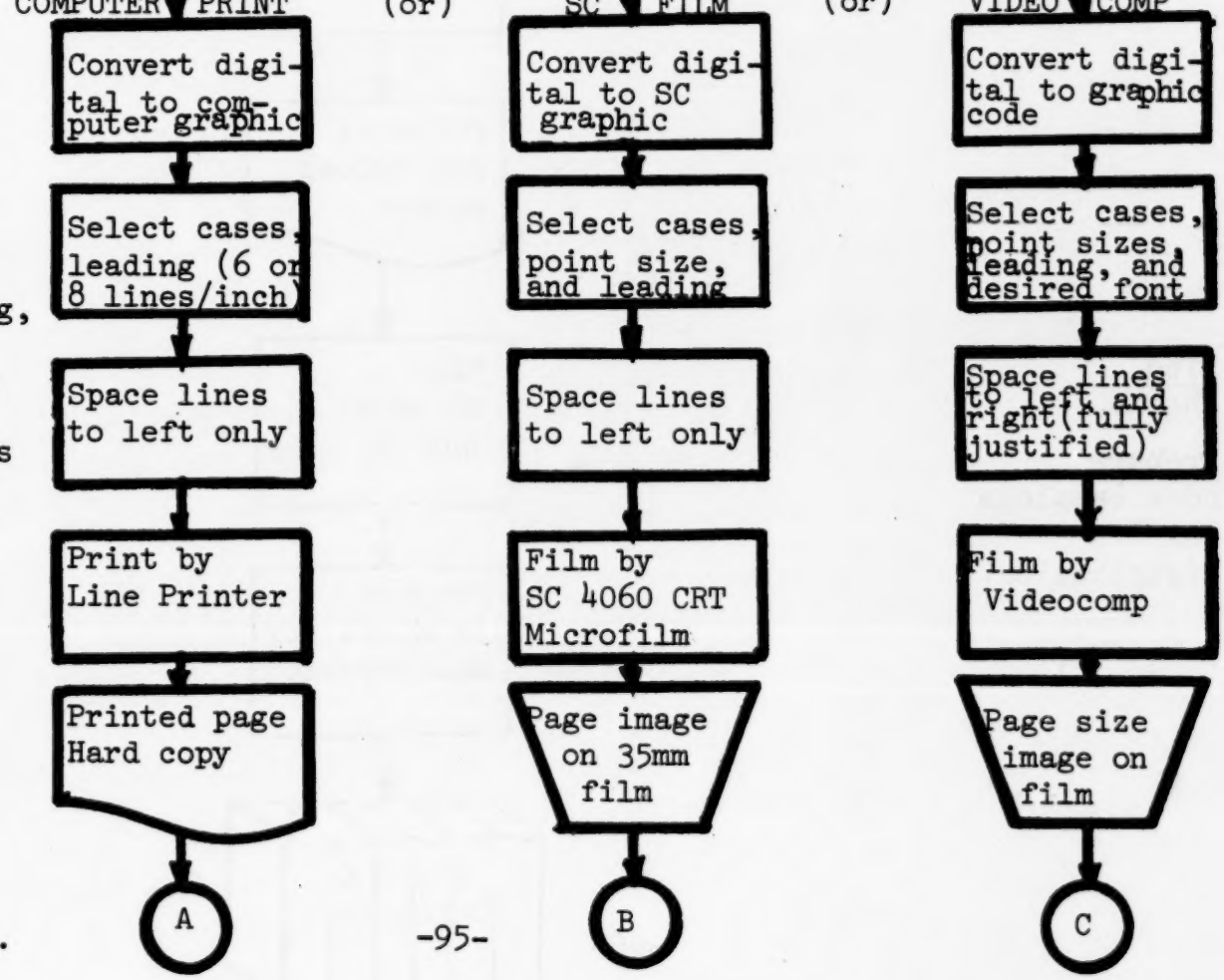
2. Convert digital codes on tape to graphics

   To format page:

   * Select case, point size, leading, (and font)

   Space lines to left or both margins

   Produce page image.


COMPUTER PRINT
(or)
SC FILM
(or)
VIDEO COMP
Convert digital to computer graphic
Select cases, leading (6 or 8 lines/inch)
Space lines to left only
Print by Line Printer
Printed page Hard copy
A
Convert digital to SC graphic
Select cases, point size, and leading
Space lines to left only
Film by SC 4060 CRT Microfilm
Page image on 35mm film
B
Convert digital to graphic code
Select cases, point sizes, leading, and desired font
Space lines to left and right(fully justified)
Film by Videocomp
Page size image on film
C

*Different weight(bold) included at extra expense.

STEP

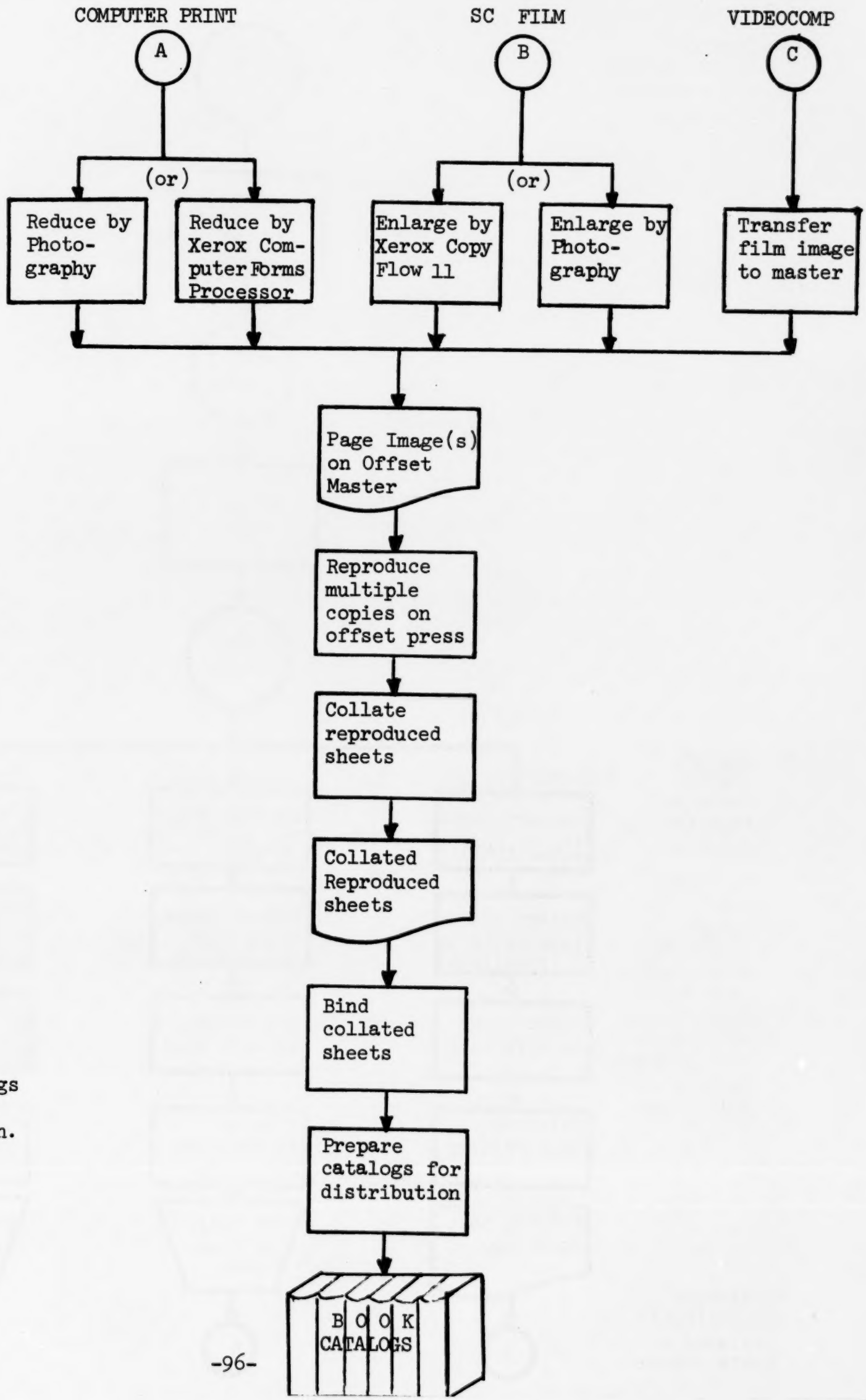
3. Create offset master (paper).

4. Reproduce multiple copies

   Collate sheets ready for binding.

5. Bind sheets

   Prepare book catalogs for distribution.



IV. OUTPUT SERVICES - COST FACTORS

(Additional Information Relating to Preceding Flowcharts)

Step 1:

The time for generation of added entries involves reading the input tape record (which includes directory and subfield delimiters in MARC form) and then writing the main and added entries on tape. For a tape pass rate of 40,000 characters/sec., we assume that the entry generation process is I/O bound - that is, compute time is less than input/output time. Thus, costs for the added entry generation may be determined by considering only the magnetic tape pass rate.

$COST Added Entry Creation = [computer cost/hr] x [Avg. # char/title (MARC form) x [1 + $\sum_{types}$ Avg. # type entries/title] + $\sum_{types}$ Avg. # char/type sort field] x [tape read/write rate (ms/char)]/[60K ms/min x 60 min/hr]

$COST Sorting = [computer cost/hr] x $\sum_{types}$ # type records [(record sort time per type record + pass time per type record pass) x 1 pass + (record merge time per type record + pass time per type record pass) x # passes] x 1/(60K ms per min x 60 min per hr)

Steps 2 and 3:

Implicit in the digital to graphic conversion is the specification of page format. For the analysis presented in this report the following figures apply:

Catalog page size	8 1/2" x 11"
Page height of output print	9.5 inches
Page width of output print	6.9 inches
Number of columns/page	2
Inter column gap	0.6 inches
Lines/inch, char/inch ratio	0.6 inches
Print density	13, 15, 18, 20 char/inches

Of interest to the user is the number of book catalog pages for a given number of unique titles.

pages/# titles = [Avg. # characters/# titles]/[# characters/page]

where,

Avg. # char/title = [(Avg. # Author Added Entries/Title) x (proportion Full Entry/Author Added Entry)

+ (Avg. # Subject Entries/Title) x (proportion Full Entry/Subject Entry)

+ (Avg. # Title Entries/Title) x (proportion Full Entry/Title Entry)

+ (Avg. # Series Entries/Title) x (proportion Full Entry/Series Entry)

+ (Main Entry)] x (Avg. # char/Full Entry)

= [(0.4)(0.6) + (1.4)(1.0) + (1.0)(0.6) + (0.05)(1.0) + 1.0] x (225)

= 740.25 char/title

and,

characters/page = (# char/inch)^2 x [lines/inch, char/inch ratio x page height of output print] x [page width of output print - inter column gap x (# columns/page - 1)]

Having the # pages/copy we are in position to compute the first page creation cost/copy.

$COST First Page Creation/Copy per Char. Density = # pages per copy x [$cost digital to graphic conversion per page per char density + $cost offset master creation per page]

Three different ways of achieving a page image have been explored in this analysis, but several different options are available for producing the offset master from these page images.

SC1 represents the cost estimates from George's Litho (San Francisco) who would carry the entire production process from prepared magnetic tapes to bound catalog copies.

SC2 and SC3 both involve a similar procedure as SC1 except that we have subcontracted to RAND Corporation (Santa Monica) the digital to graphic conversion on the 4060; to Bay Microfilm (Palo Alto), the creation of the offset master from film negative (SC2 using the Xerox

Copyflo 11 and SC3 using photographic enlargement); and to the University's Central Duplicating Service (Berkeley), the offset printing and binding task. SC2 and SC3 represent a 60% and 40% savings, respectively, in first page creation costs as compared to SC1. The SC1 and SC3 methods yield a higher quality output character image compared with SC2 because of the photographically produced offset master. The offset master creation process in the SC2 method (Xerox Copyflo 11) is much faster, however, with a plate production rate of 3000/hr compared to an estimated 300/hr for SC1 and 30/hr for SC3. Quality controls in all three methods can be achieved by in house surveillance of SC1 or by supervision of the separate steps for the other two methods.

SV (Sedgewick Videocomp) carries the process from tape to offset master (first page creation). Reproduction is assumed to be at rates comparable to Central Duplicating.

CXL (Xerox's Computer Forms Processor) involves only the process from line printer output to offset master, but in the cost analysis we have included the line printer cost (upper and lower case font) to obtain the first page creation cost - CXU is the same method except that it involves only upper case line printer font. CPL and CPU (photo offset) also use line printer listing (lower and upper case or upper case font only) as source to the process, but a photographic reduction process (Central Duplicating, Berkeley) is used to create the offset master.

The costs involved in the digital to graphic conversion and creation of the offset master for the eight alternate methods are listed in the following table.

METHOD	DIGITAL TO GRAPHIC	MASTER OFFSET
SC1	0.35/Frame	0.42/Page
SC2	0.05/Frame	0.25/Page
SC3	0.05/Frame	0.42/Page
SV	100 + .03 Line	0.25/Page*
CXL	100 x 20K Lines/hr	0.17/Page
CXU	100 x 60K Lines/hr	0.17/Page
CPL	100 x 20K Lines/hr	0.40/Page
CPU	100 x 60K Lines/hr	0.40/Page

*Actual quoted hardcopy price is $0.15. An additional $0.07 was added to this for offset master paper cost.

Step 4:

Reproduction costs/page are independent of method selected in steps 2 and 3, however, the printing cost/copy varies with the number of copies because the first page creation cost is distributed over the number of copies.

$COST Press Run per Page = Cost labor, ink, overhead per sheet side + cost paper per sheet side + gathering price per sheet side + folding price per sheet side per fold + cutting price per sheet side per cut

= .003 + .001 + .001 + .0005 + .0002

= .0057

$COST Printing per # Copies = # copies x # pages per copy x [($cost offset master creation per page)/# copies + $cost press run per page]/# pages per sheet side

Step 5:

The binding and final preparation costs listed in the Summary were for 300 sheets/volume. The entire cost table (quoted by George's Litho) presented here includes perfect binding, cover printing, and carton packed copies.

Sheets per Volume \ No. of Copies	100	200	300	400	500
100	58.09	89.14	115.55	138.04	161.03
200	64.80	100.12	128.36	150.24	173.23
300	68.89	106.70	136.04	157.56	180.53

SELECTED BIBLIOGRAPHY

Berg, David. Cost Analysis of Alternative Methods of Producing Book Catalogs. Berkeley, University of California, 1967. (unpublished thesis)

Cartwright, Kelley L.; Shoffner, Ralph M. Catalogs in Book Form: A Research Study of Their Implications for the California State Library and the California Union Catalog, with a Design for Their Implementation. Berkeley, Institute of Library Research, University of California. January 1967.

Chapin, R. E.; Pretzer, D. H. "Comparative Costs of Converting Shelf List Records to Machine Readable Form," in Journal of Library Automation. Vol. 1, No. 1, March 1968. pp. 66-74.

Fasana, Paul J. "Determining the Cost of Library Automation" in American Library Association Bulletin, 61 (June 1967) pp. 656-661.

Hammer, Donald P. "Problems in the Conversion of Bibliographical Data - A Keypunching Experiment," in American Documentation. January 1968, Vol. 19, No. 1, pp. 12-17.

Hayes, R. M.; Shoffner, R. M. The Economics of Book Catalog Production; A Study Prepared for Stanford University Libraries and the Council on Library Resources. Sherman Oaks, California, Hughes Dynamics, Inc. Advanced Information Systems Division, May 31, 1964.

Johnson, R. D. "A Book Catalog At Stanford," in Journal of Library Automation. Vol. 1, No. 1, March 1968. pp. 13-50.

Los Angeles County Public Library, Los Angeles. An Optical Character Recognition Research and Demonstration Project. Los Angeles, Los Angeles County Public Library, 1968.

Phillips, Arthur H. Computer Peripherals and Typesetting; A Study of the Man-Machine Interface Incorporating a Survey of Computer Peripherals and Typographic Composing Equipment. London, Her Majesty's Stationery Office, 1968.

Sherman, Don; Shoffner, Ralph M. California State Library: Processing Center Design and Specifications, 3 vols. Berkeley, Institute of Library Research, University of California, April 1969.

Simmons, P. A. "An Analysis of Bibliographic Data Conversion Costs," in Library Resources of Technical Services. Vol. 12, No. 3, Summer 1968, pp. 296-311.